

Freie Universität Berlin  
Fachbereich Erziehungswissenschaft und Psychologie

**Inklusion und Bildungsstandards – gegensätzliche  
Anforderungen für Unterricht und  
Leistungsbeurteilung?**

Empirische Studien zur Perspektive von Lehrkräften

Dissertation  
zur Erlangung des akademischen Grades  
Doktorin der Philosophie (Dr. phil.)

vorgelegt von  
Katharina Holder M.A.

Berlin, 2017

Erstgutachterin: Prof. Dr. Ursula Kessels  
Zweitgutachter: PD Dr. Holger Gärtner  
Disputation: 19. April 2018

## **Danksagung**

Die Danksagung ist aus Gründen des Datenschutzes nicht enthalten.

# Inhaltsverzeichnis

<b>Zusammenfassung</b> .....	<b>6</b>
<b>Summary</b> .....	<b>9</b>
<b>Einleitung</b> .....	<b>11</b>
<b>1 Inklusion und Bildungsstandards – bedeutsame Rahmenbedingungen von Schule und Unterricht</b> .....	<b>14</b>
1.1 Inklusion .....	14
1.1.1 Ausgangspunkt und Begriffsbestimmungen .....	14
1.1.2 Umsetzung schulischer Inklusion .....	18
1.2 Bildungsstandards.....	21
1.2.1 Ausgangspunkt und Begriffsbestimmungen .....	21
1.2.2 Implementation durch die KMK .....	23
<b>2 Die Vereinbarkeit von Inklusion und Bildungsstandards</b> .....	<b>28</b>
2.1 Unterrichtsgestaltung .....	29
2.1.1 Inklusiver und standardorientierter Unterricht: theoretische Grundlagen .....	31
2.1.2 (Un)Vereinbare Implikationen von Inklusion und Bildungsstandards für Unterrichtsgestaltung? .....	39
2.2 Leistungsbeurteilung .....	44
2.2.1 Leistungen und Leistungsbeurteilung aus der Perspektive von Inklusion und Bildungsstandards.....	44
2.2.2 (Un)Vereinbare Implikationen von Inklusion und Bildungsstandards für Leistungsbeurteilung?.....	48
2.2.3 Konsequenzen für die Leistungsbeurteilung stereotypisierter Schülergruppen .....	56
<b>3 Forschungsfazit, Fragestellungen und Studienüberblick</b> .....	<b>62</b>
3.1 Forschungsfazit.....	62
3.2 Fragestellungen und Studienüberblick .....	67
<b>4 Studien</b> .....	<b>73</b>
4.1 Studie 1: Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven und standardorientierten Unterricht aus der Sicht von Lehrkräften .....	73
4.2 Studie 2: Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung.....	108

4.3 Studie 3: Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective .....	139
<b>5 Gesamtdiskussion .....</b>	<b>172</b>
5.1 Zusammenfassung zentraler Ergebnisse.....	173
5.2 Einordnung und Reflexion der Ergebnisse .....	176
5.2.1 Die Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Unterrichtsgestaltung aus der Sicht von Lehrkräften .....	177
5.2.2 Die Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Leistungsbeurteilung aus der Sicht von Lehrkräften .....	183
5.3 Limitationen.....	192
5.4 Implikationen.....	194
5.4.1 Praxis.....	195
5.4.2 Forschung.....	197
5.5 Fazit .....	201
<b>6 Literaturverzeichnis .....</b>	<b>204</b>
<b>Erklärung .....</b>	<b>225</b>
<b>Lebenslauf .....</b>	<b>226</b>

## **Zusammenfassung**

In deutschsprachigen Schulsystemen treffen aktuell die Reformansätze der Inklusion und der Bildungsstandards aufeinander, wodurch ein vielschichtiger bildungswissenschaftlicher Diskurs um deren Vereinbarkeit ausgelöst wurde. Jedoch liegt kaum Evidenz dazu vor, ob und inwiefern Lehrkräfte mit diesen aktuell bedeutsamen Reformen unterschiedliche oder gar gegensätzliche Anforderungen verbinden. Im Rahmen der vorliegenden Arbeit wurde unter Rückgriff auf sozialpsychologische und erziehungswissenschaftliche Theorien empirisch untersucht, wie sich in der Literatur diskutierte potentielle Gegensätze zwischen Inklusion und Bildungsstandards aus der Sicht von Lehrkräften darstellen. Dabei wurden Aspekte der Unterrichtsgestaltung (Studie 1) ebenso betrachtet wie Aspekte der Leistungsbeurteilung (Studie 1 und Studie 2) und Konsequenzen, die mit der Hervorhebung unterschiedlicher Beurteilungsmaßstäbe im Kontext beider Reformen für die Leistungsbeurteilung bestimmter Schülergruppen verbunden sein können (Studie 3).

Die experimentell angelegte *Studie 1* untersuchte auf Basis der theoretischen Unterscheidung von Sicht- und Tiefenstrukturen des Unterrichts (Oser & Baeriswyl, 2001) und der Theorie der Bezugsnormorientierung (BNO, Rheinberg, 1980) an einer Stichprobe von  $N = 135$  Lehrkräften im Vorbereitungsdienst, inwiefern diese mit Inklusion und Bildungsstandards eine jeweils andere Unterrichtsgestaltung und Leistungsbeurteilung verbinden. Dabei wurden zwei Perspektiven der Unterrichtsbeschreibung (Selbst- und Fremdeinschätzung) kontrastiert. Die Ergebnisse zeigten, dass sich aus Lehrersicht inklusiver Unterricht von standardorientiertem Unterricht hinsichtlich der Umsetzung zentraler Merkmale von Unterrichtsqualität und hinsichtlich der darin auftretenden BNO bei der Leistungsbeurteilung unterscheidet. Berichtete Unterschiede waren zum Teil auf Einstellungen zu und Erfahrungen mit beiden Reformen sowie diesbezügliche

Selbstwirksamkeitserwartungen seitens der Lehrkräfte zurückzuführen und variierten mit der Perspektive, aus der Unterricht beschrieben wurde.

Davon ausgehend, dass Inklusion und Bildungsstandards unterschiedliche Bezugsnormen bei der Leistungsbeurteilung (Rheinberg, 1980) implizieren, untersuchte die experimentell angelegte *Studie 2* anhand einer Stichprobe von  $N = 57$  Lehrkräften im Vorbereitungsdienst, inwiefern eine Auseinandersetzung mit Bildungsstandards oder Inklusion in Abhängigkeit von den Einstellungen zu diesen Reformen zur Anwendung unterschiedlicher Bezugsnormen führt. Die Auseinandersetzung mit Bildungsstandards zog bei Lehrkräften mit einer positiven Einstellung zu diesen eine schwächer ausgeprägte individuelle BNO sowie tendenziell eine stärker ausgeprägte soziale BNO nach sich, während diese bei Personen mit einer negativen Einstellung zu Bildungsstandards zu einer stärker ausgeprägten individuellen BNO und tendenziell zu einer schwächer ausgeprägten sozialen BNO führte. Die Auseinandersetzung mit Inklusion beeinflusste die BNO hingegen nicht, weder bei positiver noch bei negativer Einstellung zu Inklusion.

Da im Kontext beider Reformen unterschiedliche Maßstäbe der Leistungsbeurteilung hervorgehoben werden, untersuchte *Studie 3* mögliche Konsequenzen der Anwendung subjektiver und objektiver Beurteilungsskalen für die Leistungsbeurteilung stereotypisierter Schülergruppen, wobei die sozialpsychologische Shifting-Standards-Theorie (z.B. Biernat, Manis & Nelson, 1991) zugrunde gelegt wurde. Die Studie gliederte sich in zwei experimentelle Vignettenstudien, wobei Teilstudie 1 ( $N = 115$  Lehramtsstudierende) negative Stereotype gegenüber Mädchen in Mathematik fokussierte und Teilstudie 2 ( $N = 265$  Lehramtsstudierende) negative Stereotype gegenüber türkischstämmigen Schülern in Deutsch. In beiden Teilstudien wurde jeweils das Mitglied der stereotypisierten Schülergruppe bei gleicher Leistung auf der objektiven Beurteilungsskala negativer eingeschätzt als das Mitglied der nicht stereotypisierten

Schülergruppe, während sich auf der subjektiven Beurteilungsskala kein Unterschied in der Einschätzung beider Schülergruppen zeigte.

Die Ergebnisse der drei Studien weisen auf bedeutsame Aspekte hin, hinsichtlich derer aus der Sicht von Lehrkräften unterschiedliche Anforderungen mit Inklusion und Bildungsstandards verbunden sein können, und tragen somit Erkenntnisse zur wahrgenommenen Vereinbarkeit beider Reformen in Bezug auf Unterrichtsgestaltung und Leistungsbeurteilung bei. Abschließend werden aus den Ergebnissen Vorschläge für die pädagogische Praxis sowie weitere Fragestellungen für die zukünftige Forschung abgeleitet.



## Summary

At present, two main reforms are implemented simultaneously within the educational systems of German speaking countries: inclusion and educational standards. The issue of whether the two are compatible figures prominently in current scientific debate. Although it is discussed in the literature that these reforms may have contradictory implications for teaching and performance evaluation, only few empirical studies have addressed this issue from a teachers' perspective. The present dissertation attempts to fill this gap using approaches from educational sciences and social psychology. For that purpose, aspects of instructional quality (Study 1) as well as aspects of performance evaluation (Study 1 and Study 2) and consequences were examined, which may occur for specific social groups of students since both reforms highlight different approaches to performance evaluation (Study 3).

The experimental *Study 1* was based on the theoretical distinction between surface structures and deep structures of instruction (Oser & Baeriswyl, 2001) and the theory of reference norm orientation (Rheinberg, 1980). It examined how trainee teachers ( $N = 135$ ) perceive instructional quality and performance evaluation in the context of inclusion and educational standards respectively, taking account of two perspectives on teaching – the own perspective and the external perspective. Trainee teachers associated different forms of instructional quality and different reference norm orientations with inclusion and educational standards. Attitudes towards both reforms, experiences with them and self-efficacy beliefs partly explained these differences. Further, reported differences between the two reforms varied depending on the perspective from which instructional quality and performance evaluation were described.

Given inclusion and educational standards are associated with different reference norm orientations (Rheinberg, 1980) and taking account of subjects' preexisting attitudes,

the experimental *Study 2* tested whether exposure to the concepts of “inclusion” vs. “educational standards” affected trainee teachers’ ( $N = 57$ ) reference norm orientations. Teachers preferred either an individual reference norm or a social reference norm depending on their attitudes towards educational standards after being exposed to this concept. Confronting teachers with inclusion had no effect on their reference norm orientations regardless of a priori attitudes towards inclusion.

Inclusion and educational standards highlight different approaches to performance evaluation. Against this background and based on the shifting standards theory (e.g., Biernat, Manis & Nelson, 1991) *Study 3* examined possible consequences of applying subjective or objective scales on performance evaluation of negatively stereotyped students. The study was divided in two experimental vignette studies. Substudy 1 ( $N = 115$  student teachers) focused on the stereotype that boys are better at mathematics than girls, substudy 2 ( $N = 265$  student teachers) addressed the stereotype that students with a Turkish immigrant background perform worse at school than nonimmigrant students, especially in the subject German. Both substudies revealed that stereotyped students were evaluated less favorably than nonstereotyped students when objective scales were used, but were evaluated similarly to nonstereotyped students when subjective scales were applied.

In summary, the results point to important aspects for which inclusion and educational standards may have conflicting implications from a teachers’ perspective. Thus, they contribute relevant findings to the current scientific debate on possibly conflicting demands imposed on teachers by inclusion vs. standardization reforms. Finally, implications for future research and practice are suggested based on the findings.

## **Einleitung**

Das Handeln von Lehrkräften ist durch verschiedene Spannungsfelder geprägt, die sich aus bestimmten, für den Lehrerberuf grundlegenden Antinomien ergeben können (z.B. Helsper, 2002). Spannungsverhältnisse, die in der Literatur diskutiert werden, beziehen sich z.B. auf die gleichzeitig erforderliche Orientierung an gesellschaftlichen Zielvorstellungen und der Individualität jeder Schülerin und jedes Schülers oder auf die gleichzeitigen Ansprüche des Förderns und Selektierens, welche an Schule herangetragen werden (z.B. Bräu, 2008). Zwei Reformansätze, die in deutschsprachigen Schulsystemen aktuell aufeinandertreffen und mit bedeutsamen Implikationen für Schule und Unterricht verbunden sind, verleihen diesen grundlegenden Spannungsverhältnissen des Lehrerhandelns eine neue Brisanz: die zunehmende Umsetzung schulischer Inklusion einerseits und die Implementation bundesweit geltender Bildungsstandards andererseits. Während mit Inklusion eine starke Konzentration auf die individuell unterschiedlichen Lernvoraussetzungen und -möglichkeiten innerhalb einer immer heterogeneren Schülerschaft einhergeht, werden mit den Bildungsstandards einheitliche Lernziele für alle Schülerinnen und Schüler formuliert, an deren Erreichen Schul- und Unterrichtsqualität gemessen werden soll.

Inklusion und Bildungsstandards bestimmen den aktuellen bildungswissenschaftlichen Diskurs um Schule und Unterricht maßgeblich (Eversheim, 2015). Es wird darauf hingewiesen, dass die beiden Reformansätze einander hinsichtlich unterschiedlicher Aspekte, die sowohl Unterrichtsgestaltung als auch Leistungsbeurteilung betreffen, konträr gegenüberstehen können (z.B. Biewer, 2012; Seitz, 2006b; Werning & Arndt, 2015). Die Vereinbarkeit dieser aktuell bedeutsamen Reformen stellt somit eine vielschichtig diskutierte Problematik dar. Demgegenüber existieren jedoch kaum empirische Erkenntnisse, inwiefern Lehrkräfte, die als betroffene Akteure beide Reformen umsetzen müssen, unterschiedliche oder gar widersprüchliche Anforderungen mit diesen verbinden. Fundierte Erkenntnisse zur

Sichtweise von Lehrkräften sind jedoch unerlässlich, um einerseits den Diskurs um die Vereinbarkeit von Inklusion und Bildungsstandards differenziert und begründet führen zu können und andererseits Herausforderungen und damit einhergehende Unterstützungsbedarfe zu identifizieren, die sich eventuell aus dem Aufeinandertreffen beider Reformen ergeben.

Ziel der vorliegenden Arbeit ist es, zu überprüfen, ob und inwiefern sich die in der Literatur diskutierte potentielle Gegensätzlichkeit zwischen Inklusion und Bildungsstandards aus der Sicht von Lehrkräften empirisch bestätigen lässt. Dabei werden Aspekte der Unterrichtsgestaltung ebenso betrachtet wie Aspekte der Leistungsbeurteilung und Konsequenzen, die mit der Hervorhebung unterschiedlicher Beurteilungsmaßstäbe im Kontext beider Reformen für die Leistungsbeurteilung bestimmter Schülergruppen verbunden sein können. Den Kern der Arbeit bilden drei empirische Studien, denen eine gemeinsame theoretische Rahmung vorausgeht und an die sich eine gemeinsame Diskussion der gefundenen Ergebnisse anschließt.

In Kapitel 1 des Theorieteils werden Inklusion und Bildungsstandards einführend als bedeutsame Rahmenbedingungen von Schule und Unterricht vorgestellt. Kapitel 2 beleuchtet anschließend die Vereinbarkeit beider Reformen in Bezug auf Unterrichtsgestaltung und Leistungsbeurteilung unter Berücksichtigung des diesbezüglichen bildungswissenschaftlichen Diskurses. Unter Rückgriff auf die theoretische Unterscheidung von Sicht- und Tiefenstrukturen des Unterrichts (Oser & Baeriswyl, 2001), die Theorie der Bezugsnormorientierung (BNO, Rheinberg, 1980) und die Shifting-Standards-Theorie (z.B. Biernat, Manis & Nelson, 1991) wird aufgezeigt, welche potentiellen Spannungsverhältnisse, die das Handeln von Lehrkräften betreffen, sich zwischen Inklusion und Bildungsstandards identifizieren lassen. In Kapitel 3 wird darauf aufbauend zunächst ein Forschungsfazit gezogen. Weiterhin werden die sich daraus ableitenden Fragestellungen vorgestellt und es wird ein Überblick über die Ziele und die

Methodik der drei empirischen Studien, die zu deren Überprüfung konzipiert und durchgeführt wurden, gegeben. Das darauffolgende Kapitel 4 setzt sich aus den drei Studien zusammen, die jeweils in Manuskriptform dargestellt werden. Es schließt sich mit Kapitel 5 eine Gesamtdiskussion der drei Studien an, die zunächst deren zentrale Ergebnisse zusammenfasst. Anschließend werden die Ergebnisse und Limitationen vor dem Hintergrund bisheriger Forschungsergebnisse und aktueller bildungswissenschaftlicher Diskurslinien diskutiert. Darauf aufbauend werden Vorschläge für die schulische Praxis entwickelt und Perspektiven für weitere Forschung aufgezeigt, bevor die Arbeit mit einem studienübergreifenden Fazit schließt.

## **1 Inklusion und Bildungsstandards – bedeutsame Rahmenbedingungen von Schule und Unterricht**

Die zunehmende Umsetzung von Inklusion einerseits und die verbindliche Einführung von Bildungsstandards andererseits stellen im deutschsprachigen Bildungssystem bedeutsame Rahmenbedingungen für Schule und Unterricht dar. Im folgenden Kapitel werden die grundlegenden Charakteristika beider Reformansätze vorgestellt, wobei deren Ausgangspunkte, zugehörige zentrale Begriffe und Konzepte sowie deren jeweilige Umsetzung in Deutschland dargestellt werden.

### **1.1 Inklusion**

#### **1.1.1 Ausgangspunkt und Begriffsbestimmungen**

Mit Inkrafttreten der UN-Konvention über die Rechte von Menschen mit Behinderungen (UN-BRK, United Nations, 2006) im Jahr 2009 verpflichtete sich Deutschland zur Schaffung eines inklusiven Bildungssystems (BGBl. 2008, Teil II, Nr. 35). Vorausgegangen war bereits im Jahr 1994 die Konferenz "Pädagogik für besondere Bedürfnisse: Zugang und Qualität" der Organisation der Vereinten Nationen für Bildung, Wissenschaft und Kultur (UNESCO) aus der die sog. Salamanca-Erklärung (UNESCO, 1994) hervorging. Diese wurden neben Deutschland von 91 weiteren Staaten unterzeichnet und betonte bereits, dass „jedes Kind ein grundsätzliches Recht auf Bildung hat“ (UNESCO, 1994, o. S.) und dass auch Schülerinnen und Schülern mit besonderen Bedürfnissen der Zugang zu allgemeinen Schulen ermöglicht werden muss (UNESCO, 1994). Erst durch die Ratifizierung der UN-BRK wurde jedoch die verbindliche, rechtliche Grundlage für die gesellschaftliche Inklusion von Menschen mit Behinderungen geschaffen (BGBl. 2008, Teil II, Nr. 35). Der für das Bildungssystem entscheidende Artikel 24 der UN-BRK trifft grundlegend folgende Regelung: „Die Vertragsstaaten anerkennen das Recht von Menschen mit Behinderungen

auf Bildung. Um dieses Recht ohne Diskriminierung und auf der Grundlage der Chancengleichheit zu verwirklichen, gewährleisten die Vertragsstaaten ein integratives<sup>1</sup> Bildungssystem auf allen Ebenen“ (BGBl. 2008, Teil II, Nr. 35, S. 1436). Alle Länder der Bundesrepublik Deutschland waren in der Folge dazu aufgefordert, das Recht auf inklusive Bildung in ihren jeweiligen Schulgesetzen zu verankern (Lange, 2017).

Infolge des Inkrafttretens der UN-BRK intensivierte sich auch die bildungspolitische und -wissenschaftliche Debatte um schulische Inklusion (Hinz, 2013; Werning, 2014). Innerhalb dieser Debatte fallen die Definitionen und Verständnisse von Inklusion sehr unterschiedlich aus und fokussieren unterschiedliche Bedeutungsebenen (Göransson & Niholm, 2014; Grosche, 2015; Löser & Werning, 2015). Diese Verschiedenheit aufgreifend, wurden mehrere Systematiken zur Strukturierung der verschiedenen Inklusionsverständnisse vorgeschlagen. Ainscow, Booth und Dyson (2006a, S. 15) fassen beispielsweise sechs Verständnisse von Inklusion zusammen:

- Inklusion bezogen auf Schülerinnen und Schüler mit Einschränkungen oder Behinderungen, beschrieben mit der Kategorie „sonderpädagogischer Förderbedarf“,
- Inklusion als Antwort auf Ausgrenzung von Schülerinnen und Schülern aus disziplinarischen Gründen,
- Inklusion bezogen auf alle Schülergruppen, die von Ausgrenzung betroffen sind,
- Inklusion als Entwicklung einer „Schule für alle“,
- Inklusion in Bezug auf Bildung für alle sowie
- Inklusion als grundlegende Herangehensweise an Gesellschaft und Bildung.

---

<sup>1</sup> In der amtlichen deutschen Fassung wurde das englische Wort „inclusive“ mit „integrativ“ übersetzt. Rechtlich verbindlich sind nur die Fassungen des Übereinkommens in den sechs UN-Sprachen (Englisch, Französisch, Spanisch, Russisch, Chinesisch, Arabisch) (Beauftragte der Bundesregierung für die Belange von Menschen mit Behinderungen, 2017)

Göransson und Niholm (2014) kamen hingegen in ihrem internationalen Forschungsüberblick zu dem Schluss, dass die meisten der empirischen Forschung zugrundeliegenden Begriffsverständnisse von Inklusion vier verschiedenen Kategorien zugeordnet werden können. Inklusion wird entweder verstanden als

- Platzierung von Schülerinnen und Schülern mit Behinderungen in Regelklassen,
- Aufgreifen der sozialen und akademischen Bedürfnisse von Schülerinnen und Schülern mit Behinderungen,
- Aufgreifen der sozialen und akademischen Bedürfnisse aller Schülerinnen und Schüler oder
- Schaffung von Lerngemeinschaften mit spezifischen Charakteristika (Göransson & Niholm, 2014, S. 268).

In diesen beiden Strukturierungsversuchen der verschiedenen international vorliegenden Verständnisse von Inklusion deutet sich an, dass sich grundlegend zwei unterschiedliche Perspektiven unterscheiden lassen (Ainscow et al., 2006a; Löser & Werning, 2015; Ruberg & Porsch, 2017). Einerseits ein “special educational needs’ view of inclusion” (Ainscow et al., 2006a, S. 16), der auf die Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf bzw. Behinderungen fokussiert, wie die folgende Definition der Ständigen Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (KMK) exemplarisch veranschaulicht: „Grundlage inklusiver Bildung sind das gemeinsame Lernen und die gemeinsame Erziehung von Kindern und Jugendlichen mit und ohne Behinderungen“ (KMK, 2011, S. 7). Andererseits wird von anderen Autorinnen und Autoren ein eher weites Inklusionsverständnis vertreten, das verschiedenen Dimensionen von Heterogenität in Schülergruppen Rechnung trägt (Ainscow et al., 2006a; Löser & Werning, 2015; Ruberg & Porsch, 2017). Als Beispiel dafür kann folgende Definition von Hinz (2010, S. 192) angeführt werden: “Inklusion bemüht sich, alle Dimensionen von



Heterogenität in den Blick zu bekommen und gemeinsam zu betrachten. Dabei kann es um unterschiedliche Fähigkeiten, [...] ethnische Herkunft, Nationalitäten, [...] soziale Milieus, [...] körperliche Bedingungen oder anderes mehr gehen“.

Verschiedene Autorinnen und Autoren kamen vor diesem Hintergrund zu dem Schluss, dass Inklusion ein mehrdimensionales, uneinheitlich definiertes Konstrukt mit „konzeptuellen Unschärfen“ (Grosche, 2015, S. 20) und, wie Artiles, Kozleski, Dorn und Christensen (2006, S. 69) beschreiben, eine „broad, boundary-blurring agenda“ darstelle. Im deutschsprachigen Raum kommt außerdem die Besonderheit hinzu, dass eine klare Unterscheidung von Inklusion und Integration nur bedingt gelingen kann, da beide Begriffe teilweise synonym und teilweise in Abgrenzung zueinander verwendet werden (Grosche, 2015; Hinz, 2002; Kocaj, Kuhl, Kroth, Pant & Stanat, 2014). In der empirischen Bildungsforschung ist das am meisten untersuchte Konstrukt im Kontext von Inklusion jedoch die gemeinsame Beschulung von Schülerinnen und Schülern mit und ohne sonderpädagogischen Förderbedarf (Göransson & Niholm, 2014; Grosche, 2015; Lindsay, 2007; Ruijs & Peetsma, 2009). Als Gemeinsamkeit der verschiedenen dargestellten Definitionen kann zudem herausgestellt werden, dass mit schulischer Inklusion stets eine Vergrößerung der Heterogenität innerhalb der Schülerschaft verbunden ist.

Zusammenfassend lassen sich unterschiedliche Perspektiven auf bzw. Verständnisse von Inklusion identifizieren. Dieser Arbeit wird ein eher engeres Verständnis von Inklusion im Sinne einer gemeinsamen Beschulung von Schülerinnen und Schülern mit und ohne sonderpädagogischen Förderbedarf zugrunde gelegt. Da diese gemeinsame Beschulung das in der empirischen Bildungsforschung am meisten untersuchte Konstrukt darstellt, ist eine Anschlussfähigkeit eigener an frühere Forschungsergebnisse gegeben. Dennoch wird anerkannt, dass schulische Inklusion über die Thematik der Platzierung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Regelklassen hinausgeht (Löser &

Werning, 2015). Nachfolgend wird die Umsetzung schulischer Inklusion in Deutschland beschrieben.

### **1.1.2 Umsetzung schulischer Inklusion**

Aus dem oben genannten, dieser Arbeit zugrundeliegenden Begriffsverständnis leitet sich ab, dass der sog. Inklusionsanteil den zentralen Indikator für die Umsetzung inklusiver Beschulung darstellt. Dieser beschreibt den prozentualen Anteil jener Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, die eine allgemeinbildende Schule besuchen, an allen Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In einem europäischen Vergleich von 24 Staaten der European Agency for Development in Special Needs Education (EASIE) lag der Inklusionsanteil in Deutschland mit 31,5% im Jahr 2014 deutlich unter dem europäischen Gesamtdurchschnitt von 52,7% (EASIE, 2017). Anzumerken ist jedoch, dass ein internationaler Vergleich solcher Kennzahlen beispielsweise aufgrund der unterschiedlichen statistischen Erfassung und Definition von Lernenden mit sonderpädagogischem Förderbedarf in den jeweiligen Staaten nur begrenzt aussagekräftig ist (Brüggemann & Tegge, 2016).

Detaillierte Statistiken über die Umsetzung schulischer Inklusion und deren Entwicklung in Deutschland stammen von der KMK und dem Statistischen Bundesamt und wurden in einem aktuellen Überblick von Lange (2017) zusammengefasst. Darin zeigt sich für das Schuljahr 2015/2016, dass 37,7% der insgesamt 517.384 Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf eine allgemeinbildende Schule besuchten, während es im Schuljahr 2008/2009 noch 18,4% waren (KMK, 2016b; Lange, 2017). Diesem gestiegenen Inklusionsanteil steht jedoch eine nur geringfügig gesunkene Exklusionsquote gegenüber, die den Anteil von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf, die eine Förderschule besuchen, an der Gesamtschülerzahl beschreibt. Diese betrug 4,9% im Schuljahr 2008/2009 und 4,4% im Schuljahr 2015/2016 (Lange, 2017). Weiterhin

ist ein Anstieg des Anteils von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an der Gesamtschülerzahl zu verzeichnen. Diese sog. Förderquote stieg von 6,0% im Schuljahr 2008/2009 auf 7,1% im Schuljahr 2015/2016 (Lange, 2017) an. Der gestiegene Inklusionsanteil bei einer gleichzeitig nahezu konstanten Exklusionsquote kann folglich dadurch erklärt werden, dass anteilig bei mehr Schülerinnen und Schülern ein sonderpädagogischer Förderbedarf festgestellt wurde (Klemm, 2015).

Betrachtet man die Umsetzung inklusiver Beschulung in den einzelnen Ländern der Bundesrepublik Deutschland, zeigt sich ein sehr heterogenes Bild (Blanck, 2015; KMK, 2016b). Die Inklusionsanteile der Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf<sup>2</sup> im Schuljahr 2015/2016 lagen zwischen 25,4% im Land Hessen und 83,5% im Land Bremen (Lange, 2017). Auch die Exklusionsquoten variierten erheblich im Bereich von 1,1% im Land Bremen bis 6,5% im Land Mecklenburg-Vorpommern (Lange, 2017). Die Daten der empirischen Studien dieser Arbeit stammen aus den Ländern Berlin und Brandenburg, wo die Inklusionsanteile im Schuljahr 2015/2016 bei 60,4% (Berlin) und 47,5% (Brandenburg) und damit über dem Bundesdurchschnitt und die Exklusionsquoten bei 3,0% (Berlin) und 4,3% (Brandenburg) leicht unter dem Bundesdurchschnitt lagen (Lange, 2017). Blanck (2015) stellte zudem in einem alle Bundesländer vergleichenden Überblick fest, dass die qualitative Umsetzung schulischer Inklusion von einem durchgängig gemeinsamen Unterricht von Schülerinnen und Schülern mit und ohne sonderpädagogischen Förderbedarf bis hin zu gemeinsamen Aktivitäten, die ausschließlich im außerunterrichtlichen Bereich angesiedelt sind, reicht. Dies kann darauf zurückgeführt werden, dass die bildungspolitischen Strategien und Bemühungen zur Umsetzung von Inklusion in den einzelnen Ländern stark variieren (für einen Überblick siehe Lange, 2017).

---

<sup>2</sup> In manchen Bundesländern wie z.B. Niedersachsen und Nordrhein-Westfalen wird der Begriff „sonderpädagogischer Unterstützungsbedarf“ verwendet.

Zu beachten ist im Kontext gemeinsamer Beschulung, dass nicht für alle Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf an Regelschulen auch die Rahmenlehrpläne dieser Schulen gelten. In Deutschland werden, abgesehen von länderspezifischen Besonderheiten, allgemein die sonderpädagogischen Förderschwerpunkte „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“, „geistige Entwicklung“, „körperliche und motorische Entwicklung“, „Sehen“, „Hören“ und „langandauernde Krankheit“ unterschieden (KMK, 1994, S. 6f.). Schülerinnen und Schüler, die einen Förderbedarf in den Förderschwerpunkten „Lernen“ und „geistige Entwicklung“ aufweisen, werden zieldifferent, also nach eigens für diese Förderschwerpunkte geltenden Rahmenlehrplänen unterrichtet<sup>3</sup>, während Schülerinnen und Schüler mit einem Förderbedarf in anderen Förderschwerpunkten zielgleich, also nach dem allgemeinen Rahmenlehrplan der jeweiligen Schulform, unterrichtet werden (Blanck, 2015). An allgemeinen Schulen machten Schülerinnen und Schüler mit den Förderschwerpunkten „Lernen“ (ca. 46%), „emotionale und soziale Entwicklung“ (ca. 25%) und „Sprache“ (ca. 13%) im Schuljahr 2016/2017 die größten Anteile an allen Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf aus (Statistisches Bundesamt, 2017).

Zusammengefasst stieg der Inklusionsanteil von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Deutschland zwar in den letzten Jahren an, jedoch ging die Exklusionsquote gleichzeitig nur leicht zurück. Darüber hinaus werden zwischen den Bundesländern bezüglich der Umsetzung schulischer Inklusion deutliche Unterschiede sichtbar. Nichtsdestotrotz besucht inzwischen mehr als jedes dritte Kind mit sonderpädagogischem Förderbedarf eine Regelschule, womit in Einklang mit den verschiedenen Konzepten von Inklusion eine Steigerung der Heterogenität der Schülerschaft verbunden ist. Bevor

---

<sup>3</sup> In welchen Schulformen bzw. Schularten zieldifferenten Unterricht ermöglicht wird, unterscheidet sich ebenfalls zwischen den Ländern der Bundesrepublik Deutschland (für einen Überblick siehe Lange, 2017)

beschrieben wird, welche Implikationen dies für die praktische Tätigkeit von Lehrkräften haben kann, wird die Einführung der Bildungsstandards als zweite bestimmende Reform im deutschen Bildungssystem näher beleuchtet.

## **1.2 Bildungsstandards**

### **1.2.1 Ausgangspunkt und Begriffsbestimmungen**

Mit dem sogenannten „Konstanzer Beschluss“ legte die KMK im Jahr 1997 fest, dass das deutsche Schulsystem im Rahmen wissenschaftlicher Untersuchungen international verglichen werden soll (KMK, 2004a). Infolgedessen beteiligt sich Deutschland an den internationalen Leistungsvergleichsstudien „Programme for International Student Assessment“ (PISA) seit 2000, „Trends in International Mathematics and Science Study“ (TIMSS) seit 1997 und „Progress in International Reading Literacy Study“/„Internationale Grundschul-Leseuntersuchung“ (PIRLS/IGLU) seit 2001 (KMK, 2016a). Aufgrund des insbesondere anfänglich unbefriedigenden Abschneidens Deutschlands in diesen Studien, vor allem gekennzeichnet durch den sog. „PISA-Schock“, wurde ein Reformprozess angestoßen, innerhalb dessen Qualitätssicherung und -entwicklung im Bildungssystem durch eine Konzentration auf die Ergebnisse (*Outputs*) von Bildungsprozessen gesteuert werden sollen (KMK, 2004a).

Kernelement dieses Reformprozesses, mit dem ein Wechsel von der bisher praktizierten Inputsteuerung (z.B. durch Lehrpläne) zu einer Outputsteuerung im Bildungssystem einherging (Altrichter & Maag Merki, 2010), bildete die Entwicklung und Einführung von bundesweit einheitlichen Bildungsstandards (KMK, 2004a; Pant, 2013). Sie stellen einen Schwerpunkt der „Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring“ dar, die von der KMK im Jahr 2006 verabschiedet und im Jahr 2016 überarbeitet wurde (KMK, 2016a). Neben der Implementation von Bildungsstandards werden darin im Sinne

der Outputorientierung die bereits angesprochene regelmäßige Teilnahme an internationalen Schulleistungsvergleichen, Verfahren zur Qualitätssicherung auf Schulebene sowie eine gemeinsame Bildungsberichterstattung von Bund und Ländern geregelt (KMK, 2016a).

Grundlage für die Entwicklung der Bildungsstandards bildete die „Expertise zur Entwicklung nationaler Bildungsstandards“ (Klieme et al., 2003), in der die Konzeption und Funktionen von Bildungsstandards charakterisiert werden. Demnach werden mit den Bildungsstandards allgemeine Bildungsziele aufgegriffen und in konkrete Anforderungen an schulisches Lernen übersetzt (Klieme et al., 2003). Somit stellen sie gemäß der Systematik, die Ravitch (1995) zur Differenzierung von Standards im Bildungssystem vorgeschlagen hat, vorrangig Leistungsstandards (*performance standards*) dar (Klieme et al., 2003). Sie legen fachspezifisch fest, welche Kompetenzen Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe erreicht haben sollen (Klieme et al., 2003). Unter Kompetenzen werden dabei in Anlehnung an Weinert (2001) die kognitiven Fähigkeiten und Fertigkeiten eines Individuums verstanden, Probleme zu lösen, sowie die damit einhergehenden Fähigkeiten und Bereitschaften, Problemlösungen in verschiedenen Situationen erfolgreich anzuwenden bzw. anwenden zu können (Klieme et al., 2003, S. 72). Zur Überprüfung des Erreichens der beschriebenen Kompetenzen sollen landes- und bundesweite Vergleichsuntersuchungen durchgeführt werden (Klieme et al., 2003).

Die Bildungsstandards, so die Expertise, sollen in Kompetenzmodellen ausdifferenziert werden, die zentrale Charakterisierungen, Entwicklungsverläufe und Abstufungen von Kompetenzen aufzeigen (Klieme et al., 2003). Darin sollen hierarchisch angeordnete Niveaustufen des Beherrschens der Kompetenzen beschrieben werden, die bei der Überprüfung der Bildungsstandards Kriterien zur Beurteilung des erreichten Ergebnisses darstellen (Klieme et al., 2003). In der Expertise wird ausdrücklich die

verbindliche Festlegung einer Niveaustufe empfohlen, hinter der keine Schülerin und kein Schüler zurückbleiben soll, ein sog. „Mindeststandard“ (Klieme et al., 2003, S. 27).

Den so charakterisierten Bildungsstandards werden nach Klieme et al. (2003) grundlegend zwei Funktionen zugeschrieben: eine Orientierungs- und eine Diagnosefunktion. Einerseits sollen sie verbindliche Ziele darstellen, auf die sich Lehrkräfte bei der Schul- und Unterrichtsentwicklung beziehen können (Orientierungsfunktion). Andererseits sollen auf Basis der Bildungsstandards Lernergebnisse in standardbasierten Tests erfasst und beurteilt werden, um zu überprüfen inwiefern erwünschte Kompetenzen tatsächlich erreicht wurden (Diagnosefunktion) (Klieme et al., 2003). Damit soll einerseits ein Beitrag zum Bildungsmonitoring und zur Schulevaluation geleistet werden und andererseits sollen Hinweise für individuelle Förderung und Diagnostik bereitgestellt werden (Klieme et al., 2003). Von einem Einbezug der Ergebnisse standardbasierter Tests in die Notengebung und Zertifikatvergabe wird jedoch ausdrücklich abgeraten (Klieme et al., 2003).

Zusammenfassend werden mit den Bildungsstandards erwünschte Lernergebnisse (in Form von Kompetenzen) für alle Schülerinnen und Schüler bundesweit einheitlich festgelegt. Diese sollen in Kompetenzstufenmodellen konkretisiert und ihr Erreichen mit standardisierten Tests überprüft werden. Sie sollen Orientierungspunkte für Schul- und Unterrichtsentwicklung sowie eine Basis zur Diagnose von Lern- und Leistungsständen bieten. Im folgenden Kapitel wird die Umsetzung der Bildungsstandards in Deutschland im Sinne der Ausgestaltung und Implementation durch die KMK dargestellt.

### **1.2.2 Implementation durch die KMK**

Bundesweit geltende Bildungsstandards wurden von der KMK in den Jahren 2003, 2004 und 2012 für folgende Fächer und Schulstufen bzw. Abschlüsse verabschiedet:

- Bildungsstandards für den Primarbereich in den Fächern Deutsch (KMK, 2004f) und Mathematik (KMK, 2004h) (verbindliche Einführung zu Schuljahresbeginn 2005/2006)

- Bildungsstandards für den Hauptschulabschluss in den Fächern Deutsch (KMK, 2004e), Mathematik (KMK, 2004g) und der ersten Fremdsprache (Englisch und Französisch) (KMK, 2004b) (verbindliche Einführung zu Schuljahresbeginn 2005/2006)
- Bildungsstandards für den Mittleren Schulabschluss in den Fächern Deutsch (KMK, 2003b), Mathematik (KMK, 2003c) und der ersten Fremdsprache (Englisch und Französisch) (KMK, 2003a) sowie für die Fächer Biologie (KMK, 2004c), Chemie (KMK, 2004d) und Physik (KMK, 2004i) (verbindliche Einführung zu Schuljahresbeginn 2004/2005 bzw. 2005/2006)
- Bildungsstandards für die Allgemeine Hochschulreife in den Fächern Deutsch (KMK, 2012b), Mathematik (KMK, 2012c) sowie der fortgeführten Fremdsprache (Englisch und Französisch) (KMK, 2012a) (gültig für Abiturprüfungen ab dem Schuljahr 2016/2017)

Diese Bildungsstandards der KMK konzentrieren sich der Expertise von Klieme et al. (2003) folgend auf die Kernbereiche der jeweiligen Fächer und formulieren erwartete Lernergebnisse (KMK, 2004a). Somit stellen sie vorrangig Leistungsstandards dar, vermischt mit Inhaltsstandards (*content standards*) (KMK, 2004a; Ravitch, 1995). Ihr Erreichen wird, wie in der Expertise ebenfalls vorgesehen, in national einheitlichen standardbasierten Tests überprüft (KMK, 2004a). Abweichend von der Expertise beziehen sich die Bildungsstandards der KMK jedoch auf das durchschnittlich erwartete Leistungsniveau von Schülerinnen und Schülern (KMK, 2004a) am Ende der jeweiligen Schulstufe bzw. beim Erreichen des jeweiligen Abschlusses. Sie stellen demzufolge sog. „Regelstandards“ (KMK, 2004a), statt wie empfohlen Mindeststandards (Klieme et al. 2003), dar.

Auf Basis der KMK-Bildungsstandards wurden vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) in Kooperation mit Expertinnen und Experten aus



der schulischen Praxis fach- und jahrgangsspezifische Kompetenzstufenmodelle entwickelt, mit denen abgebildet und überprüft werden kann, welche Kompetenzen eine Schülerin oder ein Schüler erreicht hat (z.B. KMK, 2013). Diese werden anhand einer kontinuierlichen Kompetenzskala in Form von Punktwerten abgebildet, welche einen Mittelwert von 500 Punkten und eine Standardabweichung von 100 Punkten aufweist und durch cut-off-Werte in Kompetenzstufen unterteilt ist (z.B. KMK, 2013). Unterschieden werden auf dieser Messskala der Bildungsstandards fünf Kompetenzstufen von „Stufe I (unter Mindeststandard)“ bis „Stufe V (Optimalstandard)“, wobei gezeigte Leistungen auf „Stufe III (Regelstandard)“ den Erwartungen entsprechen, die von der KMK als Regelstandards festgelegt wurden (z.B. KMK, 2013, S. 7). Alle Kompetenzstufen enthalten Könnens-Beschreibungen, die illustrieren, welche Fähigkeiten und Fertigkeiten eine Schülerin oder ein Schüler, der die jeweilige Kompetenzstufe erreicht hat, beherrscht.

Zur Implementation und Anwendung der Bildungsstandards verpflichteten sich die Länder der Bundesrepublik Deutschland in den „Vereinbarungen über Bildungsstandards“ für die jeweiligen Schulstufen bzw. Abschlüsse (KMK, 2003d, KMK, 2004j, KMK, 2004k). Implementationsmaßnahmen sollen sich laut KMK auf die Bereiche Lehrkräfteaus- und -fortbildung, Lehrplanarbeit sowie Schul- und Unterrichtsentwicklung beziehen (KMK, 2004a). Die einzelnen Implementationsbemühungen der Bundesländer können vor allem im Bereich *Lehrkräfteaus- und -fortbildung* nur schwer überblickt werden. Wenngleich kaum belastbare Daten dazu vorliegen, wird angenommen, dass die Fortbildungsangebote ähnlich wie die Lehrkräfteausbildung je nach Land und Fach sehr unterschiedlich ausfallen (Köller, 2009b). Im Bereich *Lehrplanarbeit* zeigt sich, dass sich inzwischen in allen Bundesländern die Rahmenlehrpläne der entsprechenden Fächer auf Bildungsstandards bzw. darin beschriebene Kompetenzen beziehen, jedoch in unterschiedlicher Art bzw. unterschiedlichem Ausmaß (siehe z.B. für Baden-Württemberg: Pant, 2016; für Nordrhein-

Westfalen: Ministerium für Schule und Weiterbildung, 2008; für Thüringen: Thüringer Ministerium für Bildung, Wissenschaft und Kultur, 2016). In den Ländern Berlin und Brandenburg wurde mit Beginn des Schuljahres 2017/2018 ein neuer Rahmenlehrplan unterrichtswirksam, der (in den entsprechenden Fächern) explizit an die in den Bildungsstandards beschriebenen Kompetenzen anknüpft (Senatsverwaltung für Bildung, Jugend und Familie & Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, 2015).

Ausgangspunkt für *Schul- und Unterrichtsentwicklung* sollen neben Kompetenzstufenmodellen und zugehörigen Beispielaufgaben auch die Ergebnisse standardbasierter Testungen bilden (KMK, 2010). In allen Bundesländern wurde vor diesem Hintergrund die Überprüfung der Bildungsstandards in Vergleichsarbeiten (VERA)<sup>4</sup> und zentralen Abschlussprüfungen stark vorangetrieben, was „einen zusätzlichen Wirkhebel dar[stellt], um die Innovation in die Praxis zu implementieren“ (Wacker, Rohlfis & Kramer, 2013, S. 122). VERA kommt damit eine zentrale Vermittlungsfunktion bei der Implementation fachlicher bzw. fachdidaktischer Konzepte der Bildungsstandards zu (KMK, 2012d).

Bundesweit einheitliche VERA werden jährlich in der dritten und achten Jahrgangsstufe durchgeführt<sup>5</sup> und überprüfen die Leistungen aller Schülerinnen und Schüler in verschiedenen Domänen auf Basis der Bildungsstandards (KMK, 2016a). Seit dem Schuljahr 2008/2009 beteiligen sich alle 16 Länder der Bundesrepublik Deutschland an VERA in der dritten Jahrgangsstufe (VERA 3) und seit dem Schuljahr 2007/2008 werden VERA in der achten Jahrgangsstufe (VERA 8) geschrieben (Richter, Böhme, Becker, Pant & Stanat, 2014). Obligatorisch ist in allen Ländern für allgemeinbildende Schulen die jährliche

---

<sup>4</sup> Die Vergleichsarbeiten werden in manchen Bundesländern auch als „Lernstandserhebungen“, „KERMIT“ oder „Kompetenztests“ bezeichnet (KMK, 2017).

<sup>5</sup> In manchen Bundesländern werden zusätzlich VERA in der sechsten Jahrgangsstufe geschrieben (Richter et al., 2014).

Durchführung von VERA 3 und VERA 8 in mindestens einem Fach, darüber hinaus in weiteren Fächern freiwillig (KMK, 2012d). Die Testergebnisse der Schülerinnen und Schüler werden mithilfe der beschriebenen Kompetenzpunkteskala der Bildungsstandards abgebildet und in allen Ländern an die Schulen zurückgemeldet, wobei sich die Ausgestaltung der Rückmeldungen zwischen den Ländern unterscheidet (KMK, 2016a). In Berlin und Brandenburg werden Individualrückmeldungen zum Abschneiden der einzelnen Schülerinnen und Schüler sowie klassen- und schulbezogene Rückmeldungen erstellt (Brunner et al., 2016). Damit ist die Erwartung verbunden, dass eine Auseinandersetzung mit den Testergebnissen eine datenbasierte Schul- und Unterrichtsentwicklung anstößt (KMK, 2010).

Zusammenfassend hat die KMK bundesweit geltende Bildungsstandards für einige Kernfächer und Schulstufen bzw. Abschlüsse verabschiedet, wenngleich (noch) nicht für alle. Abweichend von der Expertise zur Entwicklung nationaler Bildungsstandards (Klieme et al., 2003) wurden dabei verbindliche Regelstandards festgelegt, die sich auf durchschnittlich erwartete Leistungen von Schülerinnen und Schülern beziehen. Alle Bundesländer haben sich zur Implementation und Anwendung der Bildungsstandards verpflichtet, wobei die Überprüfung der Standards in nationalen VERA neben der Lehrplanarbeit und der Lehrkräfteaus- und -fortbildung als zentraler Implementationsmechanismus angesehen werden kann.

### **2 Die Vereinbarkeit von Inklusion und Bildungsstandards**

Inklusion und Bildungsstandards stehen im deutschen Bildungssystem als Reformen nebeneinander, die jeweils mit bedeutsamen Veränderungen für Schule und Unterricht verbunden sind. Dies führt zu der Frage, inwiefern beide Reformen in der schulischen Praxis gleichzeitig umsetzbar sind. Im bildungswissenschaftlichen Diskurs wurde diese Frage vielfach aufgegriffen (Ahrbeck, 2016; Ainscow et al., 2006a; Biewer, 2012; Brokamp, 2005; Eversheim, 2015; Moser Opitz, 2006; Schuck, 2014; Seitz & Platte, 2006; Werning & Arndt, 2015). Der Diskurs um die Vereinbarkeit von Inklusion und Bildungsstandards bzw. Individualisierung und Standardisierung betrifft dabei sowohl Aspekte der *Unterrichtsgestaltung* als auch der *Leistungsbeurteilung* und somit zentrale Aspekte des Lehrerhandelns.

Der Vergrößerung der Heterogenität der Schülerschaft um Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durch das Streben nach schulischer Inklusion soll sowohl bei der Unterrichtsgestaltung als auch bei der Leistungsbeurteilung vordergründig mit Maßnahmen der Individualisierung und Differenzierung begegnet werden (Büker & Meier, 2017; Gebhardt, Schwab, Krammer, Gasteiger-Klicpera & Sälzer, 2014; Heinzl & Prengel, 2012; Prengel, 2017; Seitz, 2006b; Teumer, 2012). Gleichzeitig soll mit Einführung der Bildungsstandards eine Orientierung an übergreifenden, einheitlichen Lernzielen für alle Schülerinnen und Schüler sowie eine regelmäßige Überprüfung von Schülerleistungen anhand objektiver Maßstäbe erfolgen (Klieme et al., 2003; KMK, 2010). Wie Buschkühle, Duncker und Oswald (2009, S. 7) zusammenfassen, bilden „Heterogenität und Standardisierung [...] in der aktuellen Diskussion ein Spannungsfeld, das in seinen Ausmaßen kaum überschätzt werden kann“. Besonderes Augenmerk wird dabei auf das Handeln von Lehrkräften gelegt, die die Anforderungen beider Reformansätze gleichzeitig umsetzen müssen (Amrhein, 2014; Teumer, 2012).

In den folgenden Kapiteln werden die Implikationen beider Reformen für Unterrichtsgestaltung und Leistungsbeurteilung beschrieben, wobei besonders der bildungswissenschaftliche Diskurs um potentiell unterschiedliche Auswirkungen und daraus resultierende Spannungsfelder aufgegriffen wird. Die unterschiedlichen Perspektiven auf Leistung und Leistungsbeurteilung werden in diesem Diskurs gesondert hervorgehoben (Ainscow et al., 2006a; Moser Opitz, 2006; Seifert & Müller-Zastrau, 2014; Seitz, 2006b; Teumer, 2012), da im Zusammenhang mit Inklusion und Bildungsstandards u.a. jeweils unterschiedliche Maßstäbe bei der Leistungsbeurteilung im Vordergrund stehen (z.B. Seifert & Müller-Zastrau, 2014). Aus früheren sozialpsychologischen Studien ist bekannt, dass die Art des gewählten Beurteilungsmaßstabs sich insbesondere auf die Beurteilung von Mitgliedern stereotypisierter<sup>6</sup> Gruppen auswirken kann (Biernat et al., 1991; Biernat & Manis, 1994; Biernat & Vescio, 2002). Aufgrund dessen werden zusätzlich die Konsequenzen des Anwendens verschiedener Beurteilungsmaßstäbe, die jeweils durch Inklusion und Bildungsstandards hervorgehoben werden, für stereotypisierte Schülergruppen beleuchtet.

## **2.1 Unterrichtsgestaltung**

Unterrichtsgestaltung und -qualität kann anhand unterschiedlicher Merkmale analysiert und beschrieben werden, wobei grundlegend Sicht- und Tiefenstrukturen von Unterricht unterschieden werden (Kunter & Trautwein, 2013; Oser & Baeriswyl, 2001). Während Sichtstrukturen sich auf beobachtbare, übergeordnete Organisationsmerkmale des Unterrichts beziehen, verweisen Tiefenstrukturen auf basale Dimensionen, welche die Interaktionsebene zwischen Lehrenden und Lernenden kennzeichnen (Kunter & Voss, 2011; Kunter & Trautwein, 2013).

---

<sup>6</sup> Die Formulierung „stereotypisiert“ wird zur Vereinfachung stets im Sinne von „*negativ* stereotypisiert“ verwendet.

Als Sichtstrukturen gelten dabei Organisationsformen, wie z.B. Klassen- der Kursunterricht, Sozialformen, wie z.B. Gruppen- oder Partnerarbeit, und Methoden, wie direkte Instruktion oder Projektarbeit (Kunter & Trautwein, 2013; Kunter & Ewald, 2016). Auf Ebene der Tiefenstrukturen besteht ein breiter Konsens, dass kognitive Aktivierung, konstruktive Unterstützung und effiziente Klassenführung zentrale Basisdimensionen von Unterrichtsqualität darstellen (Helmke, 2009; Klieme, Schümer & Knoll, 2001; Kunter & Voss, 2011). Kognitive Aktivierung beschreibt das Potenzial des Unterrichts, tiefgehende Prozesse des Nachdenkens und der aktiven Auseinandersetzung mit einem Lerngegenstand anzuregen, z.B. durch anregende Aufgabenstellungen oder Unterrichtsgespräche (Kunter & Voss, 2011). Konstruktive Unterstützung charakterisiert Aktivitäten, die auf die Herstellung einer kognitiv, motivational und emotional unterstützenden Lernumgebung zielen, wie z.B. den konstruktiven Umgang mit Fehlern oder das Bereitstellen von Hilfestellungen zur Aufgabenbearbeitung (Kunter & Voss, 2011). Effiziente Klassenführung beschreibt koordinierende und steuernde Strategien, die zum Ziel haben, dass die vorhandene Lernzeit optimal genutzt wird (Kunter & Voss, 2011).

Empirisch wurde vielfach bestätigt, dass sich Unterrichtsgestaltung bzw. Unterrichtsqualität auf das Lernen von Schülerinnen und Schülern auswirken kann (für einen Überblick siehe Seidel & Shavelson, 2007). Dabei ließ sich feststellen, dass Sichtstrukturen den Rahmen für Unterrichtsgestaltung bilden, die dahinterliegenden Tiefenstrukturen jedoch wirkmächtiger erscheinen (Seidel et al., 2006; Seidel & Shavelson, 2007), da diese die Grundlage für eine qualitätsvolle Ausgestaltung der Sichtstrukturen bilden (Klieme, 2006; Kunter & Voss, 2011). So wurde gezeigt, dass eine hohe Qualität der kognitiven Aktivierung und der Klassenführung sich vor allem positiv auf Schülerleistungen auswirkt, während sich für konstruktive Unterstützung vor allem positive Effekte auf emotionale und motivationale Merkmale nachweisen ließen (Baumert et al., 2010; Fauth, Decristan, Rieser,

Klieme & Büttner, 2014; Kunter et al., 2013; Seiz, Decristan, Kunter & Baumert, 2016). Im Folgenden wird dargestellt, welche Implikationen für Unterrichtsgestaltung sich aus den Reformen Inklusion und Bildungsstandards ergeben. Diese werden unter Rückgriff auf die Unterscheidung von Sicht- und Tiefenstrukturen beleuchtet.

### **2.1.1 Inklusiver und standardorientierter Unterricht: theoretische Grundlagen**

Die Implementation von Inklusion und Bildungsstandards zielt jeweils darauf ab, Unterricht zu verändern bzw. zu verbessern (KMK, 2010, 2011). Jedoch ergeben sich aus den Bildungsstandards selbst keine direkten Handlungshinweise für die Unterrichtsgestaltung (Wacker et al., 2013), sondern erst aus den diese ergänzenden Handreichungen (Blum, Drüke-Noe, Hartung & Köller, 2012; Drieschner, 2009; KMK, 2010) oder Leitfäden zur Nutzung der VERA für Unterrichtsentwicklung (z.B. Benzing et al., 2016; Emmrich, Ernst, Harych & Wesselhöft, 2012). Ähnliches gilt für Inklusion, wobei gegenwärtig zwar erste Konzepte für inklusive Unterrichtsgestaltung formuliert werden (Carle, 2017; Heimlich, 2007; Kullmann, Lütje-Klose & Textor, 2014; Reich, 2014; Seitz, 2006a), jedoch (noch) keine Evidenz für das Vorhandensein einer speziellen inklusiven Didaktik vorliegt (Pool Maag & Moser Opitz, 2014; Werning & Baumert, 2013). Im Folgenden werden zuerst Merkmale inklusiven Unterrichts auf Basis von in Deutschland häufig rezipierten Ansätzen zur inklusiven Unterrichtsgestaltung dargestellt. Weiterhin werden Merkmale standardorientierten Unterrichts anhand einschlägiger Handreichungen zur Unterrichtsgestaltung auf Basis der Bildungsstandards bzw. VERA beschrieben.

#### *Inklusiver Unterricht*

Als Kernelement inklusiven Unterrichts wird der konstruktive Umgang mit Heterogenität betrachtet, durch den allen Schülerinnen und Schülern mit ihren individuellen Voraussetzungen und Bedürfnissen ein erfolgreiches Lernen ermöglicht werden soll

(Heimlich, 2007; Kullmann et al., 2014; Reich, 2014; Seitz, 2006a; Werning & Arndt, 2015). Dies setzt eine Anerkennung der Individualität jedes und jeder Lernenden durch die Lehrkräfte voraus (Carle, 2017; Kullmann et al., 2014). Auf der Ebene der Sichtstrukturen kommt dabei Maßnahmen der individuellen Förderung und inneren Differenzierung eine große Bedeutung zu, durch die eine optimale Passung von den Lerninhalten und -zielen zu den heterogenen Lernvoraussetzungen der Schülerinnen und Schüler hergestellt werden soll (Kullmann et al., 2014; Seitz, 2006a; Seitz & Scheidt, 2012; Werning & Arndt, 2015). Lehrkräfte können dies auf unterschiedliche Weise realisieren (Klieme & Warwas, 2011; Trautmann & Wischer, 2007). So könnten z.B. unterschiedliche Lernerfahrungen über eine große Methodenvielfalt und offene Unterrichtsformen ermöglicht und/oder differenzierende Aufgaben(-formate) angeboten werden (Heimlich, 2007; Kullmann et al., 2014; Seitz, 2006a). Undifferenzierter Frontalunterricht erscheine hingegen für inklusive Lerngruppen weniger sinnvoll (Heimlich, 2007). Wenngleich innere Differenzierung im Vordergrund steht, sollen Lehrkräfte im inklusiven Unterricht gleichzeitig einen Fokus auf die bewusste Herstellung von Gemeinsamkeit legen, z.B. durch das Realisieren sozialer und kooperativer Lernformen und/oder die Arbeit an einem gemeinsamen Lerngegenstand (Hennies & Ritter, 2015; Kullmann et al., 2014; Seitz, 2006a; Seitz & Scheidt, 2012). Durch die Umsetzung der „pädagogische[n] Leitidee individueller Herausforderung in sozialer Eingebundenheit“ (Seitz & Scheidt, 2012, o. S.) soll verhindert werden, dass Individualisierung zu Vereinzelung führt (Hennies & Ritter, 2015).

Auf Ebene der Tiefenstrukturen wird die besondere Bedeutung einer effizienten Klassenführung für inklusive Lerngruppen hervorgehoben (Emmer & Stough, 2001; Werning & Baumert, 2013), die z.B. durch eine effektive Nutzung der Lernzeit und einen guten Umgang mit Störungen gekennzeichnet sein soll. Im Hinblick auf konstruktive Unterstützung kann festgestellt werden, dass diese Basisdimension in den beschriebenen Ansätzen



inklusive Unterrichtsgestaltung nicht gesondert betont wird (Moser Opitz, 2014), diesen jedoch inhärent ist. Das aufmerksame Begleiten aller Lernenden, welches sich in konstruktiver Unterstützung niederschlägt (Kunter & Voss, 2011), ist im Kontext der grundlegenden Anerkennung der Individualität jeder Schülerin und jedes Schülers (Carle, 2017; Kullmann et al., 2014; Seitz & Platte, 2006) unabdingbar. Auch die vielfach betonten Maßnahmen innerer Differenzierung können bei erfolgreicher Umsetzung zur Qualität konstruktiver Unterstützung beitragen. Dass inklusiver Unterricht ebenfalls Potenziale zur kognitiven Aktivierung von Lernenden bieten soll, klingt in der Betonung der Wichtigkeit, eine optimale individuelle Förderung zu gewährleisten und unterschiedliche Lernerfahrungen zu ermöglichen (Heimlich, 2007; Kullmann et al., 2014), an.

Betrachtet man die tatsächliche Umsetzung inklusiven Unterrichts, existieren im deutschsprachigen nur wenige Studien (Budde & Blasse, 2017). Für die Ebene der Sichtstrukturen zeigte eine quantitative Befragung von Lehrkräften ( $N = 104$ ) in inklusiven Modellklassen, dass ca. 75% angaben, zumindest tendenziell differenzierende Maßnahmen und Materialien anzuwenden und integrationserfahrene Lehrkräfte berichteten, vermehrt kooperative Lernformen einzusetzen (Liebers, Seifert & Kolke, 2014). Die qualitative Studie von Pool Maag und Moser Opitz (2014) zeigte, dass Lehrkräfte vielfältige Maßnahmen nannten, die sie zur Differenzierung im inklusiven Unterricht einsetzen. Auf Ebene der Tiefenstrukturen brachte die Studie von Henke, Koch und Spörer (2015) hervor, dass Schülerinnen und Schüler ( $N = 1562$ ) an inklusiven Pilotenschulen sowohl die im inklusiven Unterricht realisierte Klassenführung und kognitive Aktivierung als auch den Umgang mit Heterogenität durch die Lehrkraft als hoch bzw. gut einschätzten. Auf internationaler Ebene kamen Davis und Florian (2004, S. 6) in ihrem Forschungsüberblick zu dem Ergebnis, dass sich Unterricht für Schülerinnen und Schüler mit *special educational needs* nicht grundsätzlich vom Unterricht in Regelklassen unterscheidet. Eine Zusammenfassung von 15 Meta-

Analysen zur Effektivität verschiedener Unterrichtsmerkmale für Schülerinnen und Schüler mit Lernbeeinträchtigungen brachte hervor, dass besonders die Vermittlung und Anwendung kognitiver Strategien sowie Elemente direkter Instruktion förderlich erschienen (Therrien, Zaman & Banda, 2010). Auch eine effiziente Klassenführung wurde als bedeutsames Gestaltungsmerkmal für inklusiven Unterricht identifiziert (Emmer & Stough, 2001). Werning und Baumert (2013) kommen auf Basis der Befundlage zur Gestaltung inklusiven Unterrichts zu dem Schluss, dass dieser durch Merkmale gekennzeichnet ist, die zu einem erfolgreichen Unterricht für *alle* Lernenden beitragen und nicht nur speziell für inklusiven Unterricht bedeutsam sind, wenngleich für Schülerinnen und Schüler mit Lernbeeinträchtigungen spezielle Aspekte der Unterrichtsgestaltung als besonders förderlich herausgestellt wurden.

Wie vielfach konstatiert wird, erscheinen bei der erfolgreichen Umsetzung inklusiven Unterrichts positive Einstellungen der Lehrkräfte in Bezug auf Inklusion bzw. Integration<sup>7</sup> von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf zentral (Boban & Hinz, 2003; Gebhardt, Schwab, Nusser & Hessels, 2015; Ruberg & Porsch, 2017; Werning, 2014). Internationale Studien belegten, dass (angehende) Lehrkräfte eine neutrale bis positive Einstellung zu Inklusion aufweisen (Avramidis & Norwich, 2002; de Boer, Pijl & Minnaert, 2011), was sich auch im deutschsprachigen Raum bestätigen ließ (Bosse & Spörer, 2014; Hellmich & Görel, 2014; Kessels, Erbring & Heiermann, 2014; Ruberg & Porsch, 2017). Wie systematische Forschungsübersichten zeigten, üben jedoch zahlreiche Moderatoren einen Einfluss auf die Einstellungen zu Inklusion aus (Avramidis & Norwich, 2002; de Boer, Pijl & Minnaert, 2011; Ruberg & Porsch, 2017). Auf Seiten der Lehrkräfte

---

<sup>7</sup> Diese Begriffe werden in der hier dargestellten Literatur nebeneinander verwendet, zielen jedoch beide auf die gemeinsame Beschulung von Schülerinnen und Schülern mit und ohne sonderpädagogischen Förderbedarf ab (Avramidis & Norwich, 2002).

wurden dabei vor allem die Erfahrungen mit Inklusion oder die diesbezüglichen Selbstwirksamkeitserwartungen untersucht<sup>8</sup>, wobei übereinstimmend festgestellt wurde, dass mit hohen Selbstwirksamkeitserwartungen und einer großen Erfahrung positivere Einstellungen zu Inklusion einhergehen (Avramidis & Norwich, 2002; de Boer et al., 2011; Ruberg & Porsch, 2017).

Es wird davon ausgegangen, dass Einstellungen bzw. Überzeugungen sich als Teil der professionellen Handlungskompetenz von Lehrkräften (Baumert & Kunter, 2006) auf deren Handeln im inklusiven Unterricht auswirken (Avramidis & Norwich, 2002). Anzumerken ist jedoch, dass aktuell kaum empirische Evidenz dazu vorliegt, inwiefern die Einstellungen zu Inklusion mit tatsächlichem Unterrichtshandeln oder den Lernergebnissen von Schülerinnen und Schülern zusammenhängen (Gasterstädt & Urban, 2016; Sermier Dessementet, Benoit & Bless, 2011). Es liegen lediglich erste Hinweise vor, dass Überzeugungen von Lehrkräften zu ihrer Rolle und ihrer Verantwortung bei der Inklusion von Lernenden mit besonderem Förderbedarf positiv mit der Qualität ihrer Unterrichtsgestaltung zusammenhängen (Jordan & Stanovich, 2001; Jordan, Schwartz & McGhie-Richmond, 2009).

### *Standardorientierter Unterricht*

Mit der Implementation der Bildungsstandards ist das übergeordnete Ziel verbunden, Unterricht kompetenzorientiert zu gestalten (Blum, 2012; Drieschner, 2009; KMK, 2010; Lenski, Richter & Pant, 2015). „Jede einzelne Unterrichtsstunde [...] muss sich daran messen lassen, inwieweit sie zur Förderung und Weiterentwicklung inhaltsbezogener und allgemeiner Schüler-Kompetenzen beiträgt“ (Blum, 2012, S. 17). Dabei soll kein vollkommen neues Unterrichtskonzept etabliert, sondern allgemeinen Kriterien „guten Unterrichts“ (z.B. Helmke, 2009) entsprochen werden (Drieschner, 2009; KMK, 2010).

---

<sup>8</sup> Darüber hinaus erwiesen sich zahlreiche weitere Variablen wie z.B. die Art des sonderpädagogischen Förderbedarfs der Lernenden, der Lehramtstyp oder die Schulform als einflussreich. Für einen Überblick siehe Avramidis und Norwich (2002), de Boer et al. (2011) sowie Ruberg und Porsch (2017).

Auf Ebene der Sichtstrukturen wird betont, dass Lernangebote so aufbereitet werden sollen, dass sie möglichst passend auf die jeweiligen Voraussetzungen der Schülerinnen und Schüler abgestimmt sind (Drieschner, 2009; KMK, 2010). Lehrkräfte sollen folglich differenzierende Ansätze der Unterrichtsgestaltung und individuelle Förderung (Klieme & Warwas, 2011) wirkungsvoll umsetzen, damit die aus den Bildungsstandards abgeleiteten Lernziele von allen Lernenden erreicht werden können (Drieschner, 2009). Ansatzpunkte zur Ableitung solcher Lernziele könnten Kompetenzstufenmodelle und/oder die Ergebnisse standardbasierter Tests darstellen (Klieme et al., 2003; KMK, 2013). Weiterhin kommt Unterrichtsformen, die eine hohe Selbsttätigkeit der Schülerinnen und Schüler anregen und eine hohe Handlungsorientierung aufweisen, wie z.B. Projektunterricht oder Freiarbeit, eine besondere Bedeutung zu (Drieschner, 2009; Drüke-Noe, 2012; KMK, 2013). In diesem Kontext wird ein Perspektivwechsel vom Lehren zum Lernen und das Etablieren einer neuen Aufgabenkultur gefordert, wodurch kompetenzorientierte Schülertätigkeiten angeregt werden sollen (Blum et al., 2012; Drieschner, 2009).

Auf Ebene der Tiefenstrukturen wird besonders hervorgehoben, dass sich standardorientierter Unterricht durch ein hohes Maß an kognitiver Aktivierung auszeichnen soll, z.B. durch das Arbeiten in Projektphasen und/oder das Anbieten anregender Aufgabenstellungen (Drieschner, 2009; Drüke-Noe, 2012). Da standardorientierter Unterricht sich an allgemeinen Kriterien guten Unterrichts orientieren soll (KMK, 2010), kann davon ausgegangen werden, dass auch die Basisdimensionen konstruktive Unterstützung und effiziente Klassenführung dabei in Interaktionsprozessen zwischen Lehrkräften und Lernenden berücksichtigt und qualitativ ausgestaltet werden sollen.

Der Einfluss von Bildungsstandards auf Unterrichtsgestaltung wurde international vor allem im Kontext von *high-stakes-tests* betrachtet, mit denen je nach Abschneiden Belohnungen oder Sanktionen für Schulen bzw. Lehrkräfte verbunden sind (Köller, 2009a;

Maag Merki, 2010; Maier & Kuper, 2012). Dabei zeigten sich sowohl positive Effekte auf Unterrichtsgestaltung (Hamilton et al., 2007) als auch negative Effekte, wie z.B. eine zeitweise niedrigere Unterrichtsqualität oder eine Verengung des Curriculums (Griffith & Scharmann, 2008; Plank & Falk Condliffe, 2013; Smith & Kovacs, 2011). Diese Ergebnisse sind jedoch nur bedingt auf die Verhältnisse in deutschsprachigen Schulsystemen übertragbar, da die hiesigen standardorientierten Tests keine ernsten Konsequenzen nach sich ziehen und entsprechend nicht als *high-stakes-tests* gelten (Maier & Kuper, 2012). Forschungsergebnisse aus dem deutschsprachigen Raum weisen zusammenfassend darauf hin, dass mit der Implementation der Bildungsstandards und deren Überprüfung durch VERA zwar vereinzelt Veränderungen des Unterrichts, wie z.B. ein zunehmender Einsatz von Projektunterricht (Wacker et al., 2013) und eine zielgerichtete Veränderung von Aufgaben (Wurster & Richter, 2016) einhergehen, jedoch meist keine substantiellen Innovationen der Unterrichtsgestaltung stattfinden (Altrichter & Posch, 2007; Groß Ophoff, 2013; Nachtigall & Hellrung, 2013; Pant, Vock, Pöhlmann & Köller, 2008).

Ähnlich wie im Kontext von Inklusion gelten die Einstellungen von Lehrkräften zu Bildungsstandards als zentraler Faktor für deren gelingende Implementation, wobei für den Anstoß von Unterrichtsentwicklungsprozessen auch die Akzeptanz von VERA als bedeutsam angesehen wird (Grillitsch, 2010; Groß Ophoff, 2013; Richter et al., 2014). Beer (2006) fand für Österreich, dass Lehrkräfte den Bildungsstandards gegenüber eine ablehnende Haltung zeigten. In einer späteren Implementationsstudie in Österreich zeigten Grillitsch und Amtmann (2012), dass die Befragten den Standards mehrheitlich neutral gegenüberstanden und für ca. 13% problematische Aspekte der Bildungsstandards überwogen. In Deutschland stehen Lehrkräfte den Standards im Mittel grundsätzlich offen gegenüber, wünschen sich jedoch mehr Information und Unterstützung im Umgang damit (Pant et al., 2008). Bezogen auf die standardbasierten VERA zeigte sich in verschiedenen

Studien, dass diese überwiegend akzeptiert und im Mittel als eher nützlich wahrgenommen werden (Kühle & Peek, 2007; Maier, 2008a; Nachtigall & Hellrung, 2013). Weiterhin wurde belegt, dass Lehrkräfte (wenn überhaupt) vor allem dann ihren Unterricht im Sinne der durch die Bildungsstandards intendierten Wirkungen weiterentwickeln, wenn sie den Tests gegenüber eine hohe Akzeptanz und Nützlichkeitsersparung bzw. positive Einstellungen aufweisen (Groß Ophoff, 2013; Maier, 2008b; Nachtigall & Hellrung, 2013; Richter et al., 2014; Wurster & Richter, 2016). So zeigten z.B. Richter et al. (2014) auf Basis von Daten der IQB-Ländervergleichsstudie aus dem Jahr 2011, dass Lehrkräfte, die VERA als Instrument zur Unterrichtsentwicklung begreifen, ihren Unterricht kompetenzorientierter und differenzierter gestalten.

Zusammenfassend wird einerseits deutlich, dass im Kontext beider Reformen teilweise *unterschiedliche* Merkmale von Unterrichtsgestaltung als besonders bedeutsam herausgestellt werden. Andererseits sollen sowohl inklusiver Unterricht als auch standardorientierter Unterricht hinsichtlich der angeführten Merkmale von Unterrichtsgestaltung bzw. -qualität bestmöglich ausgestaltet werden. Besonders auffallend ist, dass in den theoretisch-konzeptionellen Ausführungen zu standardorientiertem Unterricht die Notwendigkeit einer differenzierenden Unterrichtsgestaltung ebenso hervorgehoben wird wie in jenen zu inklusivem Unterricht. Studien zur Umsetzung inklusiven Unterrichts sind bisher eher rar, wenngleich internationale Studien darauf verweisen, dass qualitativvoller inklusiver Unterricht allgemeinen Kriterien guten Unterrichts für *alle* Lernenden entspricht. Hinsichtlich der Bildungsstandards und damit verbundener standardbasierter Tests zeigte sich, dass deren Einführung zwar zu vereinzelten, jedoch nicht zu wesentlichen Veränderungen der Unterrichtsgestaltung führte. Für die Implementation beider Reformansätze scheinen zudem die diesbezüg-

lichen Einstellungen von Lehrkräften und damit zusammenhängend auch ihre Selbstwirksamkeitserwartungen und Erfahrungen eine zentrale Rolle zu spielen. Die Betrachtung zentraler Gestaltungsmerkmale inklusiven und standardorientierten Unterrichts führt nicht zwangsläufig zu dem Schluss, dass mit Inklusion und Bildungsstandards eine grundsätzlich unterschiedliche Unterrichtsgestaltung verbunden ist. Inwiefern dennoch die Vereinbarkeit beider Reformen in Bezug auf Unterricht diskutiert wird, wird nachfolgend dargestellt.

### **2.1.2 (Un)Vereinbare Implikationen von Inklusion und Bildungsstandards für Unterrichtsgestaltung?**

Die Frage, ob und inwiefern mit der zunehmenden Umsetzung von Inklusion einerseits und der Implementation der Bildungsstandards andererseits unterschiedliche oder gar widersprüchliche Implikationen für die Unterrichtsgestaltung verbunden sind, wird von verschiedenen Vertreterinnen und Vertretern der Sonder- und Inklusionspädagogik sowie der Erziehungs- und Bildungswissenschaft theoretisch aufgegriffen. Dabei werden verschiedene potentielle Spannungsverhältnisse zwischen beiden Reformen thematisiert.

Eine individualisierende bzw. differenzierende Unterrichtsgestaltung wird, wie im vorigen Abschnitt verdeutlicht, sowohl im Kontext von Inklusion als auch von Bildungsstandards als gewinnbringend betrachtet. Vertreterinnen und Vertreter der Bildungsstandards betonen dabei, dass zwar Lernziele, nicht aber die dazu führenden Lernprozesse standardisiert werden und somit Spielräume für eine individualisierte Unterrichtsgestaltung entstehen könnten (KMK, 2004a, 2010; Oelkers & Reusser, 2008). Die Vorgabe einheitlicher und verbindlicher Lernziele stehe jedoch dem Gedanken der Zieldifferenz gegenüber, den inklusive Unterrichtsgestaltung grundlegend dahingehend aufweise, dass „verschiedene Schüler/-innen unterschiedliche Ziele zur gleichen Zeit verfolgen oder aber gleiche Ziele zu unterschiedlichen Zeitpunkten“ (Werning & Arndt, 2015, S. 66). Schipolowski, Stanat und Rjosk (2017) weisen in diesem Kontext darauf hin,

dass Bildungsstandards nur für jene Schülerinnen und Schüler angelegt werden sollten, die zielgleich, also nach den Lehrplänen der allgemeinen Schulen, unterrichtet werden. Seitens der Inklusionspädagogik wird jedoch ein individualisiertes (Hinz, 2003) oder zumindest individualisierungsfähiges (Prenzel, 2015b) Curriculum für *alle* Lernenden gefordert, auch wenn derzeit nach offizieller Vorgabe nur Lernende mit bestimmten Förderschwerpunkten zielfähig unterrichtet werden (siehe Kapitel 1.1.2) (Werning & Arndt, 2015). Vor dem Hintergrund, dass für inklusiven Unterricht auch die bewusste Herstellung von Gemeinsamkeit zentral erscheint, sieht Brokamp (2005) jedoch durch die Vorgabe gemeinsamer Lernziele auch eine Chance in den Bildungsstandards, eine gemeinsame Orientierung an allgemeinen Bildungszielen für den Unterricht zu schaffen.

Bezogen auf Art und Ausmaß einer differenzierenden Unterrichtsgestaltung werden weiterhin die Kompetenzstufenmodelle der Bildungsstandards und ihre Anwendbarkeit aus inklusionspädagogischer Sicht diskutiert. Kompetenzstufenmodelle werden vor allem im Kontext standardbasierter Tests zur Verortung von Schülerleistungen herangezogen (KMK, 2010). Die Kompetenzstufen werden dabei durch Könnens-Beschreibungen illustriert, wodurch Stärken und Schwächen der Lernenden, z.B. auf Basis ihres Abschneidens in VERA, aufgezeigt werden können (Pant, 2013). Dadurch würde der Ausgangspunkt für individuelle Förderung und Differenzierung im Unterricht geschaffen (Klieme & Leutner, 2006; KMK, 2010). Seitens der Sonder- und Inklusionspädagogik wird jedoch postuliert, dass eine vertikal strukturierte Ausdifferenzierung von Kompetenzniveaus „nur schwer mit einer inklusiven Didaktik zu vereinbaren“ (Seitz, 2006b, S. 195) ist. Inklusiver Unterricht baut auf der Annahme auf, dass Lernen als individualisierter, diskontinuierlicher Prozess verläuft, dessen hohe Komplexität mit Kompetenzstufenmodellen nicht abbildbar ist (Seitz, 2006b). Aufgrund ihrer zu geringen Differenzierung und ihrem linearen Aufbau würden



sich diese kaum als Referenzsystem eignen, Lernprozesse im inklusiven Unterricht differenziert und unterstützend zu begleiten (Seitz, 2006b; Seitz & Platte, 2006). Kompetenzstufenmodelle stellten folglich keine geeignete Grundlage dar, den heterogenen Bedürfnissen und Voraussetzungen der Schülerinnen und Schüler im inklusiven Unterricht zu begegnen, sondern drückten eher den Wunsch aus, Lernprozesse zu kontrollieren und zu homogenisieren (Seitz & Platte, 2006).

Darüber hinaus wird thematisiert, dass es für inklusive Unterrichtsgestaltung von entscheidender Bedeutung ist, *wie* die Bildungsstandards durch die KMK definiert werden (Boban & Hinz, 2013; Eversheim, 2015; Geiling & Gille, 2005; Seitz, 2006b; Werning & Arndt, 2015). Kritisiert wird die Vorgabe von Regelstandards, auf deren Erreichung Unterricht ausgerichtet werden soll, da dies mit einer Orientierung an „Durchschnittskindern“ einhergehen und zur Diskriminierung und Marginalisierung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf führen kann (Biewer, 2012; Eversheim, 2015; Moser Opitz, 2006; Ritter & Hennies, 2015; Seitz, 2006b). Einige Autorinnen und Autoren sehen im Kontext inklusiver Beschulung daher die Festlegung von Mindeststandards<sup>9</sup> als geeigneter an, diese Schülerinnen und Schüler stärker einzubeziehen (Eversheim, 2015; Geiling & Gille, 2005; Werning, 2013). Andere vertreten den Standpunkt, dass sowohl Mindest- als auch Regelstandards eine Diskriminierung von Schülergruppen, die diese nicht erreichen, nach sich ziehen und daher als Orientierung für die Gestaltung inklusiven Unterrichts ungeeignet erscheinen (Biewer, 2012; Sander, 2005; Seitz, 2006b).

Zudem zeigt sich, dass auch allgemein befürchtete negative Begleiterscheinungen der Bildungsstandards im Diskurs um deren Vereinbarkeit mit Inklusion in der Literatur

---

<sup>9</sup> Beispielsweise wurden in der Schweiz mit dem sog. HarmoS-Konkordat Bildungsstandards als Mindeststandards formuliert (Schweizerische Konferenz der kantonalen Erziehungsdirektoren, 2007).

aufgegriffen werden, wie z.B. eine inhaltliche Reduzierung des Curriculums, eine Ökonomisierung von Bildungsprozessen oder eine Verbreitung von *teaching to the test* (z.B. Brügelmann, 2015; Klein, 2010; Spinner, 2005). Diese gelten generell als unerwünscht, werden jedoch für inklusiven Unterricht als besonders hinderlich angesehen (Biewer, 2012; Moser Opitz, 2006; Seitz & Platte, 2006).

Der dargestellte Diskurs um eventuell unterschiedliche Implikationen von Inklusion und Bildungsstandards für Unterrichtsgestaltung fällt folglich im deutschsprachigen Raum relativ breit aus. Es liegen jedoch kaum empirische Erkenntnisse dazu vor, inwiefern Lehrkräfte als betroffene Akteure beide Reformen generell als miteinander vereinbar ansehen (Hofmann, 2011) oder unterschiedliche Implikationen für Unterrichtsgestaltung mit beiden Reformen verbinden. Qualitative Forschungsergebnisse aus England (Ainscow, Booth & Dyson, 2006b; Dyson, Gallannaugh & Millward, 2003) deuten an, dass Lehrkräfte durchaus Spannungen zwischen Inklusion und der dort implementierten Standard-Agenda wahrnehmen und deren parallele Umsetzung kritisch beurteilen. Diese sind jedoch kaum auf den deutschsprachigen Kontext übertragbar. Für diesen existieren nur wenige Hinweise, ob und inwiefern sich Individualisierung und Standardisierung aus Sicht von Lehrkräften konträr gegenüberstehen können. In einer Befragung österreichischer Lehrkräfte ( $N = 866$ ) zur Rezeption der Bildungsstandards fand Beer (2006), dass 88% der Befragten keine Möglichkeit zur größeren Individualisierung im Unterricht mithilfe der Standards sahen. Kühn und Racherbäumer (2013) untersuchten das Spannungsfeld von Individualisierung und Standardisierung im Kontext zentraler Abschlussprüfungen. Sie zeigten an einer Stichprobe von  $N = 501$  nordrhein-westfälischen Lehrkräften, dass diese eine Unvereinbarkeit individueller Förderung und zentraler, standardbasierter Prüfungen nur teilweise wahrnahmen (Kühn & Racherbäumer, 2013). Weiterhin fanden die Autorinnen in einer ähnlichen Studie jedoch, dass die Umsetzung einer differenzierenden

Unterrichtsgestaltung durch den Druck zentraler Prüfungen eingeschränkt erlebt wurde (Racherbäumer & Kühn, 2013). Sie resümierten, dass sich standardisierte Rahmenbedingungen und individuelle Förderung aus Sicht von Lehrkräften nicht per se ausschließen, jedoch teilweise nicht intendierte negative Effekte durch Standardisierungsprozesse hervortreten können (Racherbäumer & Kühn, 2013).

Zusammenfassend wird im Diskurs um unterschiedliche bzw. ggf. unvereinbare Implikationen von Inklusion und Bildungsstandards für Unterrichtsgestaltung die Frage aufgegriffen, inwieweit sich Bildungsstandards als Bezugspunkte für die Planung und Durchführung inklusiven Unterrichts eignen. Vor allem wird thematisiert, inwiefern im Rahmen der Standardisierung von Lernzielen die für inklusiven Unterricht wesentliche innere Differenzierung angemessen umgesetzt werden kann, wenngleich diese auch in theoretischen Konzeptionen standardorientierten Unterrichts als bedeutsam herausgestellt wird. Empirisch liegen vereinzelt Hinweise vor, dass Standardisierungsprozesse aus Sicht von Lehrkräften eine differenzierende Unterrichtsgestaltung einschränken können. Gesicherte Erkenntnisse, inwiefern Lehrkräfte je eine andere Ausgestaltung von Unterrichtsmerkmalen auf Ebene der Sicht- und Tiefenstrukturen von Unterricht mit Inklusion und Bildungsstandards verbinden, fehlen jedoch.

### 2.2 Leistungsbeurteilung

Neben den eventuell gegensätzlichen Auswirkungen auf Unterrichtsgestaltung werden im Diskurs um die Vereinbarkeit von Inklusion und Bildungsstandards auch und besonders die damit verbundenen, unterschiedlichen Perspektiven auf Leistung und Leistungsbeurteilung aufgegriffen: “On the face of it, inclusion and the standards agenda are in conflict because they imply [...] different ways of thinking about achievements“ (Ainscow et al., 2006a, S. 12). Bevor auf unterschiedliche Implikationen beider Reformen für die Leistungsbeurteilung und damit verbundene Konsequenzen für die Leistungsbeurteilung stereotypisierter Schülergruppen eingegangen wird, werden zunächst grundlegend die Perspektiven von Inklusion und Bildungsstandards auf Leistung und deren Beurteilung vorgestellt.

#### 2.2.1 Leistungen und Leistungsbeurteilung aus der Perspektive von Inklusion und Bildungsstandards

##### *Inklusion*

Zunächst ist voranzustellen, dass aktuell kein einheitlicher inklusiver Leistungsbegriff vorliegt, weshalb vielfach auf die Notwendigkeit der Entwicklung eines solchen hingewiesen wird (Seitz, 2006b; Textor, Grüter, Schiermeyer-Reichl & Streese, 2017; von Bargen, 2017). Auf Basis der Anerkennung aller Schülerinnen und Schüler in ihrer Individualität und des Wertschätzens von Heterogenität, scheint es jedoch grundlegend zentral, Leistungen stärker aus der Perspektive der jeweiligen individuellen Schülervoraussetzungen zu betrachten (Meier, Haut & Ruin, 2016; Teumer, 2012) und ein „diskriminierungskritisches Leistungsverständnis“ (Prenzel, 2017, S. 14) aufzubauen. In der Folge müsse eine Vielfalt dessen entstehen, was als Leistung gelten und beurteilt werden kann (Meier et al., 2016). Somit ist Inklusion grundsätzlich mit einer „Kritik und Herausforderung des schulischen Leistungsprinzips“ (Sturm, 2015, S. 25) verbunden. Kritisiert wird, dass innerhalb der vorherrschenden schulischen Leistungskultur

hauptsächlich die Fähigkeiten bzw. Fertigkeiten von Lernenden bewertet würden, während individuellen Voraussetzungen kaum Rechnung getragen wird (Streese, Schiermeyer-Reichl, Meyer, Moritz & Wenzel, 2017). Im Zuge des Überdenkens vorherrschender Leistungsbegriffe erscheinen aus diesem Grund (ähnlich wie im Kontext von Unterrichtsgestaltung) Individualisierung und Differenzierung zentral (Büker & Meier, 2017; Teumer, 2012).

Leistungsbeurteilung müsse vor diesem Hintergrund das Würdigen individueller Fortschritte ermöglichen und den Lernprozess der Schülerinnen und Schüler in den Mittelpunkt stellen (Biewer, 2012; Seifert & Müller-Zastrau, 2014; Teumer, 2012; Wilhelm, 2009). Auch die KMK (2011) betont dies in ihrem Beschluss zur inklusiven Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen: „Alle Kinder und Jugendlichen mit Behinderungen haben in einem inklusiven Unterricht Anspruch auf Würdigung ihrer individuellen Leistungs- und Entwicklungsfortschritte“ (KMK, 2011, S. 11). Dennoch hat Schule eine Selektions- und Allokationsfunktion inne, die, wie vielfach hervorgehoben wird, im Kontext von Inklusion besonders virulent werden kann und inklusive Leistungsbeurteilung zu einer Herausforderung oder gar zu einem Dilemma für Lehrkräfte avancieren lasse (Büker & Meier, 2017; Kottmann & Miller, 2014; Seifert & Müller-Zastrau, 2014; von Barga, 2017).

Studien zu der Fragestellung, inwiefern sich die Leitgedanken von Inklusion im Leistungsverständnis von Lehrkräften bzw. in der inklusiven Leistungsbeurteilung niederschlagen, sind bisher rar und die Befunde zudem heterogen. Erste qualitative Befunde deuten einerseits darauf hin, dass Lehrkräfte in Folge der Auseinandersetzung mit Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf die individuelle Leistungsentwicklung stärker in den Blick nahmen (Greiten, 2017). Andererseits zeigten Ruin und Meier

(2015) in einer Interviewstudie, dass Lehrkräfte im inklusiven Sportunterricht ein eher homogenisierendes Leistungsverständnis aufwiesen. Hinsichtlich inklusiver Leistungsbeurteilung fanden Liebers und Schmidt (2017) in einer quantitativen Befragung von Lehrkräften in inklusiven Modellklassen ( $N = 104$ ), dass 55% der Befragten angaben, individuelle Lernfortschritte systematisch zu dokumentieren und 90% alternative Verfahren der Leistungsbeurteilung, wie z.B. Portfolios, einsetzten. Moser, Schäfer und Kropp (2014) zeigten hingegen in einer Befragung von Lehrkräften ( $N = 327$ ), dass auch in inklusiven Settings traditionelle, summative Formen der Leistungsbeurteilung präferiert wurden.

### *Bildungsstandards*

Bildungsstandards formulieren Kompetenzerwartungen für Schülerinnen und Schüler auf verschiedenen, kriterial festgelegten Niveaustufen (Klieme et al., 2003; KMK, 2013) und stellen somit objektive Leistungsanforderungen dar. Sie lenken die Aufmerksamkeit auf die Lernergebnisse von Schülerinnen und Schülern, was eng mit ihrer Funktion im Kontext outputorientierter Steuerung im Bildungssystem zusammenhängt. Denn die Bildungsstandards legen fest, welche Qualitätskriterien ein Bildungssystem erfüllen soll, womit einhergeht, das Erreichen der Standards anhand von validen und reliablen Instrumenten zu überprüfen, um gesicherte Aussagen über die Qualität von Bildungsprozessen treffen zu können (Köller, 2010). Somit wird davon ausgegangen, dass Schülerleistungen in normierten standardbasierten Tests objektiv gemessen werden können. Sollen bzw. können Bildungsstandards vor diesem Hintergrund einen Einfluss auf Leistungsbeurteilung nehmen?

Wie in der Konzeption von Klieme et al. (2003) betont, stellen die Bildungsstandards keine Benotungs- und Prüfungsregeln dar und sollten nicht als solche behandelt werden.

Auch die KMK sieht keine Benotung von VERA vor (KMK, 2012d)<sup>10</sup>. Darüber hinaus wird von verschiedenen Autorinnen und Autoren darauf hingewiesen, dass die standardbasierten Tests keine individualdiagnostischen Aussagen anstreben, auch wenn dies von Lehrkräften oder Eltern nachgefragt wird (Klieme et al., 2003; Leutner, Fleischer, Spoden & Wirth, 2008; Pant, 2013). Vorgeschlagen wird daher, die punktuellen, standardbasierten Leistungsmessungen durch formative Beurteilungen zu ergänzen (Oelkers & Reusser, 2008; Pant, 2013).

Jedoch werden in der Expertise von Klieme et al. (2003) auch die Orientierungs- und Diagnosefunktion der Bildungsstandards und standardbezogener Tests betont. Die auf den Bildungsstandards basierenden Kompetenzstufenmodelle sollen Lehrkräften ein Referenzsystem bieten, anhand dessen sie Schülerleistungen verorten können (Klieme et al., 2003). Zudem wird konstatiert, dass Lehrkräften durch die Standards ein externer Maßstab zur Beurteilung von Schülerkompetenzen an die Hand gegeben wird, mit dessen Hilfe auch eventuell vorliegende Einseitigkeiten in der eigenen Beurteilung aufgedeckt werden können (Klieme et al., 2003). Die Bildungsstandards stellen demzufolge Kriterien zur Verortung von Schülerleistungen bereit und sollen von Lehrkräften zur Reflexion der eigenen Leistungsbeurteilung herangezogen werden, nicht jedoch zur tatsächlichen Beurteilung von Schülerleistungen. Studien zur Rezeption und Nutzung standardbasierter Leistungstests in Deutschland und der Schweiz zeigten, dass Lehrkräfte entgegen dieser Intentionen durchaus auch Informationen zur Individual- oder Selektionsdiagnose aus diesen ziehen und diese zur Ergänzung und Revidierung der eigenen Schülerbeurteilung nutzen (Kühle & Peek, 2007; Maier, 2008b, 2008a; Moser, 2003).

---

<sup>10</sup> Ungeachtet dessen kann in Mecklenburg-Vorpommern die teilweise Bewertung von VERA vorgenommen werden, wie durch die „Handreichung zur Benotung von Teilen der Vergleichsarbeiten (VERA) als schriftliche Lernerfolgskontrolle“ beschrieben wird (Institut für Qualitätsentwicklung Mecklenburg-Vorpommern, 2014).

Zusammenfassend erscheint im Kontext von Inklusion eine individuelle Perspektive auf Schülerleistungen zentral, weshalb sich auch die Leistungsbeurteilung vordergründig auf den jeweils individuellen Lernprozess und dessen Würdigung konzentrieren soll. Forschungsergebnisse zur tatsächlichen Umsetzung der Leistungsbeurteilung im inklusiven Unterricht sind bisher rar und fallen heterogen aus. Bildungsstandards nehmen hingegen eine objektive Perspektive auf Leistungen und deren Beurteilung ein, wobei die Lernergebnisse von Schülerinnen und Schülern in den Mittelpunkt gerückt werden. Wenngleich diese nicht als Benotungsmaßstäbe oder zur Individualdiagnose herangezogen werden sollen, deutet sich in verschiedenen Studienergebnissen an, dass Lehrkräfte diese bzw. die Ergebnisse standardbasierter Tests dennoch in Schülerbeurteilungen einbeziehen. Nachfolgend wird dargestellt, inwiefern möglicherweise unvereinbare Implikationen für die Leistungsbeurteilung mit Inklusion und Bildungsstandards einhergehen.

### **2.2.2 (Un)Vereinbare Implikationen von Inklusion und Bildungsstandards für Leistungsbeurteilung?**

Wie im bildungswissenschaftlichen Diskurs um unterschiedliche Implikationen von Inklusion und Bildungsstandards für Leistungsbeurteilung hervorgehoben wird, bedeutet es für Lehrkräfte „eine große Herausforderung, im [...] Spannungsfeld von Bildungsstandards und individualisierter Leistungsbewertung professionell zu agieren“ (Teumer, 2012, o. S.). Kritisiert wird seitens der Inklusionspädagogik, dass innerhalb der Bildungsstandards nur ein Ausschnitt aus einem weiten Möglichkeitsraum definiert würde, der für individuelle Leistungen als akzeptabel gilt (Walm, Schultz, Häcker & Moser, 2017). Leistungsbereiche außerhalb dieses Spektrums würden somit ausgeblendet, es komme zu einer Engführung des Leistungsbegriffs und folglich zu einer Ausgrenzung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (Ainscow et al., 2006b; Heinrich, 2010; Hörmann & Hopmann, 2011; Seitz, 2006b). Weiterhin wird hervorgehoben, dass die Fokussierung der



Bildungsstandards auf zu erreichende Ergebnisse der Idee von Inklusion widerspreche, Leistungen von den Voraussetzungen und den Potenzialen der Lernenden aus zu denken (Eversheim, 2015; Ritter & Hennies, 2015; Seitz & Platte, 2006). Die Anerkennung individueller Fortschritte gerate in den Hintergrund, da es vorrangig um das Erreichen einer Leistungsnorm, statt um die bestmögliche Förderung aller Lernenden gehe (Eversheim, 2015; Heinrich, 2010). Insofern steht die Entwicklungsorientierung von Inklusion der Ergebnisorientierung der Bildungsstandards konträr gegenüber (Biewer, 2012). Misserfolge von Schülerinnen und Schülern, die die geforderten Kompetenzen nicht erreichen, seien dabei absehbar und würden bewusst in Kauf genommen (Eversheim, 2015; Sander, 2005).

Andererseits wird angemerkt, dass die summativen Lernstandsanalysen im Kontext der Bildungsstandards eine formative individuelle Diagnostik nicht ersetzen, sondern lediglich ergänzen sollen (Pant, 2013). Seitens der Inklusionspädagogik wird zudem von manchen Autorinnen und Autoren konstatiert, dass Leistungen auch im inklusiven Unterricht nicht gänzlich losgelöst von externen Kriterien betrachtet werden sollten, wobei kriteriale Maßstäbe für inklusive Leistungsbeurteilung möglichst das gesamte Leistungsspektrum aller Schülerinnen und Schüler beschreiben sollten (Prenzel, 2015a, 2015b; Schuck, 2014). Zur möglichen Abmilderung potentieller Gegensätze zwischen Inklusion und der Vorgabe von Bildungsstandards wird weiterhin vorgeschlagen, Bildungsstandards nicht nur als *performance standards* (Ravitch, 1995) zu definieren sondern auch als sog. *opportunity-to-learn-standards*, die Rahmenbedingungen erfolgreichen Lernens beschreiben und auch zur Beschreibung der Qualität inklusiven Unterrichts herangezogen werden könnten (Eversheim, 2015; Moser Opitz, 2006).

Für Lehrkräfte ergibt sich im Spannungsfeld zwischen Entwicklungsorientierung und Ergebnisorientierung die Herausforderung, sowohl individuelle Fortschritte anzuerkennen als auch übergreifende Kriterien zur Leistungsbeurteilung heranzuziehen

(Bräu, 2008; Jantowski, 2014). Zusammenfassend postuliert Amrhein (2014, S. 18): „für den [...] Kontext eines integrativen / inklusiven Schulentwicklungsauftrags besteht eine Antinomie u.a. in der gleichzeitigen Orientierung an universalen Leistungsnormen [...] und in den durch die Inklusion formulierten Ansprüchen eines jeden Individuums, als solches im Unterricht behandelt zu werden“.

Dieses unterschiedliche Vorgehen bei der Leistungsbeurteilung lässt sich auch in der Theorie der Bezugsnormorientierung verorten (Rheinberg, 1980). Als Bezugsnorm wird nach Rheinberg (1980) der Standard verstanden, mit dem eine zu beurteilende Leistung verglichen wird. Rheinberg (1980, 1998) unterscheidet drei verschiedene Arten von Bezugsnormen. Bei der *individuellen (oder ipsativen) Bezugsnorm* wird die gezeigte Leistung einer Schülerin oder eines Schülers zu seinen oder ihren Vorleistungen in Beziehung gesetzt. Bei diesem intraindividuellen Vergleich werden Verbesserungen positiv beurteilt und Verschlechterungen negativ. Bei der *sozialen Bezugsnorm* wird die Leistung eines Lernenden zu den Leistungen einer sozialen Vergleichsgruppe, z.B. der Schulklasse, in Relation gebracht. Eine Leistung, die über dem Durchschnitt dieser sozialen Vergleichsgruppe liegt, wird positiv bewertet und eine unterdurchschnittliche Leistung negativ. Bei der *kriterialen (oder sachlichen) Bezugsnorm* wird eine Leistung mit einem Kriterium verglichen, welches zuvor extern oder intern festgelegt wurde, wie z.B. aus dem Curriculum abgeleitete Anforderungen.

Die Tendenz zur Verwendung einer dieser drei Bezugsnormen bei der Leistungsbeurteilung wird als Bezugsnormorientierung (BNO) einer Lehrkraft bezeichnet (Rheinberg, 1998). Wie Rheinberg (1998, S. 42) beschreibt, sind die drei BNO jedoch mit unterschiedlichen „blinden Flecken“ verbunden. So blendet z.B. die individuelle Bezugsnorm dauerhafte Leistungsunterschiede zwischen Lernenden aus, während die soziale Bezugsnorm Schwankungen in der Lernentwicklung einzelner Schülerinnen und

Schüler kaum erkennen lässt. Aus diesem Grund erscheint es aus der Sicht verschiedener Autorinnen und Autoren sinnvoll, unterschiedliche Bezugsnormen miteinander zu kombinieren (Gläser-Zikuda, 2010; Rheinberg, 1998; Sacher, 2009). Da die drei BNO einander jedoch tendenziell widersprüchlich gegenüberstehen, kann dies in der Praxis durchaus mit Schwierigkeiten verbunden sein (Holmeier, 2013).

Welche BNO sind vor dem Hintergrund dieser Theorie mit Inklusion und Bildungsstandards verbunden? In der inklusiven Leistungsbeurteilung stehen individuelle Förderung und Wertschätzung im Mittelpunkt, wobei das Entwicklungspotential jedes Schülers und jeder Schülerin anerkannt und „jeder Entwicklungsfortschritt [...] als Leistung bewertet“ (Wilhelm, 2009, S. 107) werden soll. Somit impliziert Inklusion die vorrangige Orientierung an der individuellen Bezugsnorm bei der Leistungsbeurteilung (Seifert & Müller-Zastrau, 2014; Seitz, 2006b; Streese et al., 2017). Mit Einführung der Bildungsstandards gewinnt hingegen die kriteriale Bezugsnorm an Bedeutung (Halbheer & Reusser, 2008; Klieme et al., 2003; KMK, 2010), denn die Bildungsstandards und die darauf basierenden Kompetenzstufenmodelle stellen Kriterien zur Verortung von Schülerleistungen bereit und konfrontieren Lehrkräfte mit einem externen Maßstab zur Einschätzung dieser (Klieme et al., 2003). Zudem wird das Erreichen der kriterial beschriebenen Kompetenzen durch VERA überprüft und an die Schulen zurückgemeldet (KMK, 2016a). In den Bildungsstandards und VERA wird somit die Chance gesehen, die Perspektive von Lehrkräften über die soziale, klasseninterne Bezugsnorm hinaus stärker auf eine kriteriale Bezugsnorm auszurichten (Halbheer & Reusser, 2008; Oelkers & Reusser, 2008). Nichtsdestotrotz heben die Ergebnisrückmeldungen der VERA neben der kriterialen auch die soziale Bezugsnorm hervor, da sie die erreichten Leistungen einer Klasse zu einer schul- oder klassenübergreifenden Bezugsgruppe in Beziehung setzen (Halbheer & Reusser, 2008; Helmke & Hosenfeld, 2003; Pant, 2013). In 15 Ländern der Bundesrepublik werden

darüber hinaus individuelle Testergebnisse der Schülerinnen und Schüler an die Schulen bzw. Lehrkräfte zurückgemeldet (Harych, Emmrich & Pant, 2017), wodurch auch ein Vergleich dieser mit den Ergebnissen der jeweiligen Klasse ermöglicht wird. In den Ergebnisrückmeldungen in Berlin werden die Leistungen einer Schülerin oder eines Schülers graphisch direkt zu den Leistungen ihrer oder seiner Klasse in Relation gebracht (Emmrich, Ernst, Harych & Meiering, 2012; Emmrich et al., 2012).

Inklusion und Bildungsstandards implizieren folglich unterschiedliche Ziele der Leistungsbeurteilung und unterschiedliche BNO (Werning, 2010): einerseits das Abbilden individueller Leistungsentwicklung durch eine vorrangige Orientierung an der individuellen Bezugsnorm und andererseits das Darstellen von Stärken und Schwächen gemessen an kriterialen Vorgaben oder dem Abschneiden einer sozialen Bezugsgruppe.

BNO von Lehrkräften können sich, wie empirisch vielfach belegt, unterschiedlich auf Schülerinnen und Schüler auswirken. Wie zahlreiche Studien zeigten, übt eine individuelle BNO der Lehrkraft insbesondere einen positiven Einfluss auf motivationale und affektive Schülermerkmale aus (für einen Überblick siehe Köller, 2005; Mischko & Rheinberg, 1995), während sich diesbezüglich eher negative Effekte der sozialen BNO feststellen ließen (Dickhäuser, Janke, Praetorius & Dresel, 2017; Trudewind & Kohne, 1982). Vor diesem Hintergrund erscheint die Hervorhebung unterschiedlicher Bezugsnormen im Kontext von Inklusion und Bildungsstandards äußerst bedeutsam. Forschungsergebnisse zu der Frage, ob und inwiefern Lehrkräfte unterschiedliche BNO bei der Leistungsbeurteilung mit beiden Reformen verbinden, liegen jedoch bisher nicht vor.

Ebenfalls liegen keine Befunde dazu vor, inwiefern Inklusion und Bildungsstandards die tatsächliche Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung beeinflussen. Dies kann jedoch vor dem Hintergrund bestehender Forschungsergebnisse zu Faktoren, die mit der BNO zusammenhängen bzw. diese beeinflussen, begründet

angenommen werden, wie nachfolgend verdeutlicht wird. Wenngleich die Befundlage zur Herkunft von BNO-Unterschieden von Lehrkräften insgesamt eher begrenzt ausfällt (Dickhäuser & Rheinberg, 2003; Köller, 2005), wurden sowohl Merkmale der Lehrkraft als auch Merkmale der Beurteilungssituation bzw. des Kontextes als Faktoren untersucht, die mit der BNO zusammenhängen oder diese beeinflussen.

Auf Ebene des Kontexts argumentiert Sacher (2009, S. 88) zunächst theoretisch, die Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung sei „untrennbar verbunden mit [...] fundamentalen erzieherischen und politisch-gesellschaftlichen Zielen“. Je nachdem, mit welchen Vorgaben sich Lehrkräfte konfrontiert sehen, können im Sinne einer Entscheidung zugunsten verschiedener Ziele von Leistungsbeurteilung andere Bezugsnormen sinnvoll erscheinen (Holmeier, 2012). Empirisch konnte Holmeier (2012, 2013) dies anhand der Einführung zentraler Abschlussprüfungen zeigen, die bei Lehrkräften im Zeitverlauf zu einer zunehmenden Orientierung an der kriterialen Bezugsnorm führte. Eine situationell variierende Anwendung unterschiedlicher Bezugsnormen je nach vorgegebenen Bewertungszielen zeigten Wilbert und Gerdes (2009). Wurde bei der Leistungsbeurteilung das Ziel der Selektion experimentell vorgegeben, griffen (angehende) Lehrkräfte ( $N = 288$ ) eher auf die soziale Bezugsnorm zurück, während mit dem Ziel der Förderung eher die Anwendung der individuellen Bezugsnorm einherging (Wilbert & Gerdes, 2009). Da Inklusion und Bildungsstandards unterschiedliche BNO hervorheben, kann angenommen werden, dass eine Auseinandersetzung mit diesen beiden aktuell bedeutsamen Reformen die Anwendung von Bezugsnormen bei der Leistungsbeurteilung ebenfalls beeinflussen könnte.

Andere Forschungsergebnisse zeigten, dass die BNO ebenso mit Merkmalen der Lehrperson selbst zusammenhängt. Dabei wurden vordergründig die individuelle und soziale BNO untersucht, selten hingegen die kriteriale BNO. So unterschieden sich

Lehrkräfte, die eine individuelle oder soziale BNO aufwiesen, in ihren Erziehungszielen (Mischo & Groeben, 1995; Mischo & Rheinberg, 1995) und ihren Zielorientierungen (Retelsdorf & Günther, 2011; Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2004), wobei eine Lernzielorientierung mit einer individuellen BNO und eine Leistungszielorientierung mit einer sozialen BNO einhergingen. Weiterhin stand die kriteriale BNO in einem positiven Zusammenhang zur Urteilsgenauigkeit von Lehrkräften (Stang & Urhahne, 2016). Mit unterschiedlichen BNO gehen seitens der Lehrkräfte außerdem verschiedene Kausalattributionen für Schulleistungen, Erwartungen an die Leistungsentwicklung von Schülerinnen und Schülern, Sanktionierungsstrategien sowie Individualisierungstendenzen im Unterricht einher (für einen Überblick siehe Köller, 2005; Rheinberg, 1980). Obwohl bisher nicht untersucht wurde, inwiefern Einstellungen von Lehrkräften deren BNO beeinflussen, zählen diese zu wichtigen Aspekten der professionellen Handlungskompetenz von Lehrenden (Baumert & Kunter, 2006) und können deshalb auch für Beurteilungshandeln bedeutsam sein. Einstellungen spielen auch bei der Implementation von Inklusion und Bildungsstandards, wie im Kontext von Unterrichtsgestaltung verdeutlicht wurde, eine zentrale Rolle (siehe Kapitel 2.1.1). Da Inklusion und Bildungsstandards mit unterschiedlichen Bezugsnormen bei der Leistungsbeurteilung in Verbindung stehen, ist anzunehmen, dass auch Einstellungen zu diesen Reformen die BNO beeinflussen könnten.

Folglich scheinen sowohl Kontexte bzw. Situationen, die unterschiedliche Ziele der Leistungsbeurteilung hervorheben als auch Merkmale der beurteilenden Lehrperson mit der BNO zusammenzuhängen bzw. diese zu beeinflussen. Aus der sozialpsychologischen Forschung ist bekannt, dass sowohl Merkmale der Situation als auch Merkmale der Person Einfluss auf gezeigtes Verhalten nehmen können (Richard, Bond & Stokes-Zoota, 2003). So wird beispielsweise angenommen, dass die Interpretation einer Situation je nach

Merkmale einer Person unterschiedlich ausfallen kann. Übertragen auf die Anwendung von Bezugsnormen ist denkbar, dass eine situationelle Auseinandersetzung mit Inklusion oder Bildungsstandards je nach Einstellungen zu diesen Reformen unterschiedliches Beurteilungsverhalten hervorrufen kann. Denn wie die sozialpsychologische Forschung vielfach zeigte, können Einstellungen sich z.B. insbesondere dann auf Verhalten auswirken, wenn diese einer Person zugänglich sind (Glasman & Albarracín, 2006), wobei die Zugänglichkeit von Einstellungen durch eine Auseinandersetzung mit dem Einstellungsobjekt hergestellt bzw. erhöht werden kann (Bargh, Chaiken, Govender & Pratto, 1992; Fazio, Chen, McDonel & Sherman, 1982; Fazio, Sanbonmatsu, Powell & Kardes, 1986). Somit ist vorstellbar, dass aufgrund der Auseinandersetzung mit Inklusion und Bildungsstandards jeweils vorhandene positive oder negative Einstellungen der Lehrkräfte zu diesen Reformen handlungswirksam werden und entweder ein mit Inklusion bzw. Bildungsstandards kongruentes oder ein dazu entgegengesetztes Beurteilungshandeln im Sinne der Anwendung unterschiedlicher Bezugsnormen nach sich ziehen.

Zusammenfassend heben Bildungsstandards die kriteriale und soziale BNO hervor, während im Kontext von Inklusion eine individuelle Bezugsnorm bei der Leistungsbeurteilung als präferiert erscheint. Dadurch deuten sich in Bezug auf die Leistungsbeurteilung mögliche gegensätzliche Anforderungen für Lehrkräfte an. Zur Wahrnehmung dieser potentiellen Gegensätze aus der Sicht von Lehrkräften oder zu deren Auswirkungen auf Beurteilungshandeln liegen bisher keine Studien vor. Jedoch ist auf Basis früherer Studien zu Faktoren, die mit der BNO zusammenhängen bzw. diese beeinflussen und aufgrund sozialpsychologischer Erkenntnisse zum gemeinsamen Einfluss von situationellen Merkmalen und Personenmerkmalen auf Verhalten, anzunehmen, dass eine Auseinandersetzung mit Bildungsstandards oder mit Inklusion (als situationeller Rahmen) je nach Einstellung zu

diesen Reformen (als Merkmal der Person) zu unterschiedlichem Beurteilungshandeln im Sinne der Anwendung unterschiedlicher Bezugsnormen führen kann.

### **2.2.3 Konsequenzen für die Leistungsbeurteilung stereotypisierter Schülergruppen**

Wie verdeutlicht wurde, implizieren Inklusion und Bildungsstandards unterschiedliche Perspektiven auf Leistung, die mit der Theorie der BNO abgebildet werden können. Im Kontext von Inklusion kommt es weniger auf die Bezugnahme auf einen übergeordneten kriterialen Maßstab oder das Bewerten von Unterschiedlichkeit an, sondern eher auf das Anerkennen individueller Fortschritte (Seitz, 2006b; Teumer, 2012; Wilhelm, 2009). Um diese angemessen feststellen und wertschätzen zu können, erscheinen über klassische Formen hinausgehende alternative oder ergänzende Formate der Leistungsbeurteilung sinnvoll (Seitz, 2006b; Textor, 2015). Objektive Formen der Leistungsbeurteilung würden hingegen im inklusiven Unterricht als „Paradoxon [erscheinen], da die unterschiedlichen Lernausgangsvoraussetzungen hier nur bedingt eine Berücksichtigung erfahren“ (Seifert & Müller-Zastrau, 2014, S. 159). Bildungsstandards stellen hingegen übergeordnete Kriterien zur Verortung von Schülerleistungen bereit, die (mit Ausnahme zieldifferent unterrichteter Lernender) für alle Schülerinnen und Schüler gleichermaßen angelegt werden können (Klieme et al., 2003). Zudem kann mit den standardbasierten VERA anhand der objektiven Messskala der Bildungsstandards (siehe Kapitel 1.2.2) ausgedrückt werden, welche Kompetenzstufe eine Schülerin oder ein Schüler erreicht hat (z.B. KMK, 2013). Die kompetenzstufenbasierten Rückmeldungen aus VERA sollen für Lehrkräfte Ansatzpunkte zur Reflexion ihrer Leistungsbeurteilung bieten (Klieme et al., 2003).

Während im Kontext von Inklusion vordergründig ein individuell subjektiver Blick auf Schülerleistungen fokussiert wird, lenken die Bildungsstandards das Hauptaugenmerk auf objektive Beurteilungsmaßstäbe. Vor diesem Hintergrund kann angenommen werden,



dass beide Reformen bei Lehrkräften das Einnehmen jeweils unterschiedlicher Perspektiven bei der Leistungsbeurteilung fördern. Besondere Relevanz kommt dieser Annahme für die Beurteilung von Schülergruppen, die mit stereotypen Annahmen belegt sind, zu, wie nachfolgend erläutert wird.

Stereotype werden als generalisierte Vorstellungen über Merkmale, Eigenschaften und/oder Verhaltensweisen von Mitgliedern einer sozialen Gruppe definiert (z.B. Hilton & von Hippel 1996, S. 240; Leyens, Yzerbyt & Schadron, 1994, S. 11). Diese können die Informationsverarbeitung in sozialen Interaktionen beeinflussen und sich auf die Eindrucksbildung auswirken (z.B. Fiske & Neuberg, 1990). Die Eindrucksbildung von Lehrkräften scheint dabei neben bestimmten Annahmen über Schülerinnen und Schüler aus niedrigeren sozialen Schichten (z.B. Lorenz, Gentrup, Kristen, Stanat & Kogan, 2016), vor allem durch zwei weit verbreitete Stereotype beeinflusst zu werden: Erstens das Geschlechterstereotyp, dass Jungen mathematisch talentierter sind als Mädchen (für einen Überblick siehe Li, 1999) und zweitens das ethnische Stereotyp, dass Schülerinnen und Schüler mit Migrationshintergrund schlechtere schulische Leistungen erbringen, als jene ohne Migrationshintergrund (für einen Überblick siehe Tenenbaum & Ruck, 2007).

In Deutschland stammen die meisten Schülerinnen und Schüler mit Migrationshintergrund aus der Türkei (Rauch, Mang, Härtig & Haag, 2016), wobei Lehrkräfte speziell an türkischstämmige Lernende niedrigere und teilweise unzutreffende Leistungserwartungen zu haben scheinen (Lorenz et al., 2016). Es liegen in Deutschland jedoch insgesamt nur wenige Studien vor, die lehrerseitige Stereotype bei der Beurteilung von Schülerinnen und Schülern beleuchten (Hirschauer & Kullmann, 2010; Lorenz et al., 2016). Einzelne Befunde aus experimentellen Studien zeigten z.B., dass (angehende) Lehrkräfte Jungen eher für mathematisch-naturwissenschaftlich orientierte Sekundarschulen empfehlen als Mädchen (Nürnberger, Nerb, Schmitz, Keller & Sütterlin, 2016) und dass türkischstämmige Lernende

von Lehrkräften schlechtere Beurteilungen für dieselbe gezeigte Leistung erhalten als jene ohne Migrationshintergrund (Glock & Krolak-Schwerdt, 2013; Sprietsma, 2013).

Ein theoretisches Erklärungsmodell, inwiefern Stereotype sich auf die Leistungsbeurteilung auswirken können, bietet die sozialpsychologische Shifting-Standards-Theorie (Biernat et al., 1991; Biernat, 2009). Auf Basis dieser Theorie lässt sich verdeutlichen, inwiefern die im Kontext von Inklusion und Bildungsstandards hervorgehobenen unterschiedlichen Beurteilungsmaßstäbe mit Konsequenzen für die Leistungsbeurteilung von stereotypisierten Schülergruppen einhergehen können.

Die Shifting-Standards-Theorie (Biernat et al., 1991; Biernat, 2009) geht grundsätzlich davon aus, dass Urteile, z.B. über die gezeigte Leistung einer Person, auf relativen Vergleichen mit den Leistungserwartungen an die soziale Gruppe basieren, der die zu beurteilende Person angehört. Stereotype über soziale Gruppen haben zur Folge, dass bei der Beurteilung von Mitgliedern dieser Gruppen unterschiedliche Standards angelegt werden (Biernat et al., 1991; Biernat, 1995). Diese stereotypenbasierten Beurteilungsstandards beruhen, so die Autorinnen und Autoren, auf Annahmen darüber, inwiefern sich soziale Gruppen in ihren Durchschnittswerten und im möglichen Wertebereich der Ausprägung eines zu beurteilenden Merkmals unterscheiden (Biernat et al., 1991; Biernat & Manis, 1994; Biernat, 2009). Beispielsweise liegen einem vorherrschenden Stereotyp zufolge sowohl der Mittelwert von Aggressivität als auch die beiden Endpunkte des Ranges bei Frauen niedriger als bei Männern (Biernat & Manis, 1994). Soll das Verhalten einer Frau hinsichtlich Aggressivität beurteilt werden, bedeutet die Einschätzung „sehr aggressiv“, dass die zu beurteilende Frau verglichen mit dem Durchschnitt und dem Range aller *Frauen* sehr aggressiv erscheint. Ein Mann mit dem selben Verhalten würde hingegen nicht als „sehr aggressiv“ eingeschätzt werden, da von *Männern* insgesamt eine höhere Aggressivität erwartet wird (Biernat & Manis, 1994). Dies

beschreibt den Mechanismus der Verschiebung des Beurteilungsstandards aufgrund von Stereotypen über soziale Gruppen, den sog. *standard shift*.

Ob und inwiefern solche stereotypenbasierten *standard shifts* bei der Beurteilung von Mitgliedern verschiedener sozialer Gruppen stattfinden, ist gemäß der Theorie abhängig von Merkmalen des Kontexts, in dem eine Beurteilung stattfindet. Es wird beispielsweise angenommen, dass Wettbewerbssituationen, wie z.B. bei der Besetzung von Führungspositionen, eher dazu führen, dass strengere Standards bei der Beurteilung von Mitgliedern stereotypisierter Gruppen angelegt werden (Biernat, 2009; Biernat, Fuegen & Kobryniewicz, 2010). Außerdem wird angenommen, dass die Art des eingesetzten Beurteilungsmaßstabs beeinflusst, ob und inwiefern es zu *standard shifts* kommt (Biernat & Manis, 1994; Biernat & Fuegen, 2001; Biernat, 2009). Dabei wird zwischen Beurteilungen auf „subjektiven Skalen“ und „objektiven Skalen“ unterschieden (Biernat & Manis, 1994, S. 5).

Es wird davon ausgegangen, dass die Einheiten *subjektiver Skalen* (z.B. Likert-Skalen von „sehr schlecht“ bis „sehr gut“) je nach urteilender und zu beurteilender Person unterschiedlich interpretiert werden können, da diese nicht kriterial verankert sind (Biernat & Manis, 1994; Biernat, 2009; Biernat et al., 2010). Sie regen, wie oben beschrieben, dazu an, gezeigte Leistungen nur mit den Erwartungen an die soziale Gruppe zu vergleichen, der die zu beurteilende Person angehört (Biernat & Manis, 1994). Ausgehend von dem Stereotyp, dass Männer besser Fußball spielen als Frauen, könnte beispielsweise mit der Einschätzung „spielt sehr gut“ ein anderes fußballerisches Können gemeint sein, je nachdem ob eine Frau oder ein Mann beurteilt wird. Somit können unterschiedliche gezeigte Leistungen auf subjektiven Skalen in denselben Einschätzungen von Mitgliedern der gemäß Stereotyp unterlegenen und Mitgliedern der gemäß Stereotyp überlegenen sozialen Gruppe

resultieren. Stereotypenbasierte Urteile werden somit auf subjektiven Skalen kaum sichtbar bzw. „maskiert“ (Biernat et al., 1991, S. 547).

*Objektive Skalen*, wie z.B. die Zentimeterskala, beruhen hingegen auf einem extern vorgegebenen Referenzrahmen und ihre Einheiten haben eine festgesetzte Bedeutung, unabhängig davon, wer urteilt oder wer beurteilt wird (Biernat et al., 1991; Biernat & Manis, 1994; Biernat, 2009). Daher wird angenommen, dass objektive Beurteilungsmaßstäbe, wie z.B. durch die Bildungsstandards bereitgestellt, dazu anregen, eine übergreifende Vergleichsperspektive bei der Einschätzung einer Person einzunehmen, also einen relativ unveränderlichen Beurteilungsstandard für alle zu beurteilenden Personen anzulegen (Biernat & Manis, 1994). Wenn beispielsweise die Leistung eines Schülers in einem standardisierten Test vorausgesagt werden soll, entspricht die Einschätzung „70 Punkte von möglichen 100 Punkten“ immer 70 Punkten, ungeachtet dessen, welcher sozialen Gruppe der Schüler angehört. Bei der Leistungsbeurteilung eines stereotypisierten Schülers auf einer solchen objektiven Skala, würde dieser schlechter beurteilt werden, als ein nicht stereotypisierter<sup>11</sup> Schüler, da durch das Einnehmen einer übergreifenden Beurteilungsperspektive seine Leistung mit den Leistungen *aller* anderen Lernenden verglichen wird und dadurch eine objektiv gute Leistung unwahrscheinlicher erscheint. Auch wenn es zunächst kontraintuitiv erscheint, werden durch Stereotype verzerrte Urteile demnach bei der Beurteilung auf objektiven Skalen eher sichtbar als auf subjektiven Skalen (Biernat & Manis, 1994; Biernat, 2009).

Folglich ergibt sich, der Shifting-Standards-Theorie zufolge, für die Leistungsbeurteilung von Angehörigen stereotypisierter Schülergruppen die Konsequenz, dass diese (gegenüber Angehörigen nicht stereotypisierter Schülergruppen) negativer beurteilt werden,

---

<sup>11</sup> Die Bezeichnung „nicht stereotypisiert“ wird zur Vereinfachung stets als Gegenteil zur Bezeichnung „stereotypisiert“ im Sinne von „nicht *negativ* stereotypisiert“ verwendet, wenngleich ggf. auch Stereotype über die so bezeichneten Schülergruppen vorliegen können.

wenn eine objektive Skala zur Beurteilung eingesetzt wird. Bei Verwendung einer subjektiven Skala, deren Ausprägungen unterschiedlich interpretiert werden können, sollte sich hingegen kein Unterschied in der Beurteilung beider Schülergruppen zeigen (Biernat & Manis, 1994; Biernat & Fuegen, 2001). Diese Annahmen konnten empirisch bereits vielfach in Bezug auf unterschiedliche Stereotype gestützt werden (Biernat & Manis, 1994; Biernat & Fuegen, 2001; Biernat, 2009; Biernat et al., 2010).

Zusammenfassend stellt die Shifting-Standards-Theorie ein empirisch vielfach gestütztes Modell zur Erklärung des Einflusses von Stereotypen auf die Leistungsbeurteilung dar und verweist auf unterschiedliche Konsequenzen des Anwendens von subjektiven und objektiven Skalen für die Leistungsbeurteilung stereotypisierter Schülergruppen. Während im Kontext von Inklusion eine individuell subjektive Perspektive auf Schülerleistungen als zentral herausgestellt wird, heben die Bildungsstandards primär objektive Beurteilungsmaßstäbe hervor. Somit können beide Reformen das Einnehmen jeweils unterschiedlicher Perspektiven bei der Leistungsbeurteilung fördern. Das zunehmende Denken in objektiven Beurteilungsmaßstäben kann, auch wenn es zunächst kontraintuitiv erscheint, gemäß der Shifting-Standards-Theorie dazu führen, dass stereotypisierte Schülerinnen und Schüler (unbeabsichtigter Weise) unzutreffend beurteilt werden. Die mit Inklusion verbundene Denkweise in individuell subjektiven Beurteilungsmaßstäben würde stereotypengeleitete Beurteilungen ebenfalls nicht verringern, ihr Auftreten wäre jedoch nach außen hin, z.B. für Schülerinnen und Schüler, kaum sichtbar. Empirische Befunde zu der Frage, inwiefern sich die jeweils durch Inklusion und Bildungsstandards hervorgehobenen Beurteilungsmaßstäbe differentiell auf die Einschätzung von stereotypisierten Schülergruppen und auswirken können, liegen jedoch bisher nicht vor.

### **3    Forschungsfazit, Fragestellungen und Studienüberblick**

Inklusion und Bildungsstandards sind jeweils mit bedeutsamen Veränderungen für Schule und Unterricht verbunden, wobei Lehrkräfte diejenigen Akteure sind, die diese auf Unterrichtsebene umsetzen sollen. Im vorigen Kapitel wurde dargestellt, welche Implikationen für die Unterrichtsgestaltung und die Leistungsbeurteilung mit der Umsetzung von Inklusion und Bildungsstandards verbunden sein können. Besonderes Augenmerk wurde dabei auf den bildungswissenschaftlichen Diskurs um deren Vereinbarkeit gelegt, den das Aufeinandertreffen beider Reformen im deutschsprachigen Raum ausgelöst hat. Zudem wurde aufgezeigt, welche Konsequenzen für die Leistungsbeurteilung stereotypisierter Schülergruppen mit der Anwendung unterschiedlicher Beurteilungsmaßstäbe, die im Kontext beider Reformen hervorgehoben werden, einhergehen können. Nachfolgend werden die daraus gewonnenen Erkenntnisse und die dabei aufgezeigten Forschungslücken in einem Forschungsfazit zusammengefasst, bevor die Fragestellungen dieser Arbeit und die empirischen Studien zu deren Untersuchung vorgestellt werden.

#### **3.1    Forschungsfazit**

Hinsichtlich der *Unterrichtsgestaltung* im Kontext von Inklusion und Bildungsstandards zeigte sich, dass teilweise jeweils unterschiedliche Gestaltungsmerkmale auf Ebene der Sicht- und Tiefenstrukturen als besonders zentral hervorgehoben werden (z.B. Drieschner, 2009; KMK, 2010; Kullmann et al., 2014; Seitz, 2006a). Jedoch sollen sowohl inklusiver Unterricht als auch standardorientierter Unterricht insgesamt eine möglichst hohe Unterrichtsqualität aufweisen. Weiterhin wird insbesondere die Notwendigkeit einer differenzierenden Unterrichtsgestaltung in den theoretischen Konzeptionen standardorientierten Unterrichts ebenso hervorgehoben wie in jenen inklusiven Unterrichts. Studien zur Umsetzung inklusiven Unterrichts in Deutschland sind bisher eher

selten (Budde & Blasse, 2017). Internationale Studien verweisen jedoch darauf, dass dieser sich durch Merkmale auszeichnet, die einen qualitätvollen Unterricht für *alle* Lernenden, nicht nur für jene mit sonderpädagogischem Förderbedarf, kennzeichnen (Davis & Florian, 2004), wobei für Schülerinnen und Schüler mit Lernbeeinträchtigungen spezielle Aspekte der Unterrichtsgestaltung als besonders bedeutsam herausgestellt wurden (Therrien et al., 2010). Hinsichtlich der Bildungsstandards und damit verbundener standardbasierter Tests zeigte sich, dass deren Einführung vereinzelt, jedoch nicht substantiell zu Veränderungen der Unterrichtsgestaltung führte (z.B. Nachtigall & Hellrung, 2013; Pant et al., 2008). Zudem scheinen für die gelingende unterrichtliche Umsetzung beider Reformen positive Einstellungen zu diesen bedeutsam zu sein (Gebhardt et al., 2015; Grillitsch, 2010; Richter et al., 2014; Ruberg & Porsch, 2017) sowie v.a. im Kontext von Inklusion die mit den Einstellungen zusammenhängenden Selbstwirksamkeitserwartungen und eigenen Erfahrungen (Avramidis & Norwich, 2002; Ruberg & Porsch, 2017).

Im dargestellten bildungswissenschaftlichen Diskurs um unterschiedliche bzw. eventuell unvereinbare Implikationen von Inklusion und Bildungsstandards für Unterrichtsgestaltung wird zusammenfassend thematisiert, ob und inwiefern sich Bildungsstandards als Referenzsystem für die Planung und Durchführung inklusiven Unterrichts eignen (z.B. Biewer, 2012; Brokamp, 2005; Seitz & Platte, 2006). Besonderes Augenmerk wird auf die Frage gelegt, inwiefern im Rahmen der Vorgabe von Bildungsstandards eine für inklusiven Unterricht angemessene innere Differenzierung möglich sei (z.B. Biewer, 2012; Brokamp, 2005; Eversheim, 2015; Seitz, 2006b). Diesem bildungswissenschaftlichen Diskurs um potentielle Gegensätze stehen jedoch kaum empirische Erkenntnisse gegenüber, inwiefern Lehrkräfte Inklusion und Standardorientierung in Bezug auf Unterrichtsgestaltung als mit-

einander vereinbar ansehen (Hofmann, 2011). Folglich ist eine offene Frage, ob und inwieweit Lehrkräfte je eine andere Unterrichtsgestaltung auf Ebene der Sicht- und Tiefenstrukturen von Unterricht mit beiden Reformen verbinden.

Hinsichtlich der *Leistungsbeurteilung* wurde zunächst dargestellt, dass im Kontext von Inklusion einer individuellen Perspektive auf Schülerleistungen eine besondere Gewichtung zukommt, weshalb Leistungsbeurteilung primär den jeweils individuellen Lernprozess einer oder eines Lernenden fokussieren soll (z.B. Seifert & Müller-Zastrau, 2014; Wilhelm, 2009). Forschungsergebnisse zur tatsächlichen Umsetzung inklusiver Leistungsbeurteilung sind jedoch bisher rar und heterogen. Für die Bildungsstandards wurde hingegen herausgestellt, dass diese eine objektive Perspektive auf Leistungen und deren Beurteilung akzentuieren und dabei die Lernergebnisse von Schülerinnen und Schülern sowie deren summative Überprüfung fokussieren (Klieme et al., 2003). Obwohl die Standards und die Ergebnisse standardbasierter VERA nicht als Benotungsmaßstäbe oder zur Individualdiagnose genutzt werden sollten (Klieme et al., 2003; KMK, 2012d), deuten einige Studien an, dass diese dennoch Eingang in die schulische Leistungsbeurteilung fanden (z.B. Kühle & Peek, 2007; Maier, 2008a).

Im bildungswissenschaftlichen Diskurs um unterschiedliche Implikationen von Inklusion und Bildungsstandards für die schulische Leistungsbeurteilung wird vor allem das Spannungsfeld thematisiert, dass Lehrkräfte sich einerseits an übergeordneten Leistungsnormen und andererseits an individuellen Entwicklungsfortschritten orientieren sollen (z.B. Amrhein, 2014; Eversheim, 2015; Teumer, 2012). Dabei wird seitens der Inklusionspädagogik die Gefahr betont, dass innerhalb der Konzentration auf erwünschte Lernergebnisse die individuellen Bedürfnisse und Voraussetzungen von Lernenden in den Hintergrund gerieten (Eversheim, 2015; Heinrich, 2010). Daran anschließend wurde herausgearbeitet, dass



beide Reformen jeweils unterschiedliche BNO (Rheinberg, 1980) bei der Leistungsbeurteilung hervorheben (Werning, 2010). Während Inklusion eine vorrangige Orientierung an der individuellen BNO impliziert (Seitz & Platte, 2006; Streese et al., 2017; Wilhelm, 2009), betonen die Bildungsstandards eher die kriteriale Bezugsnorm und, im Kontext der VERA-Rückmeldungen, auch die soziale Bezugsnorm (Emmrich et al., 2012; Helmke & Hosenfeld, 2003; Pant, 2013). Forschungsergebnisse zu der Frage, ob Lehrkräfte unterschiedliche BNO bei der Leistungsbeurteilung mit beiden Reformen verbinden, liegen jedoch bisher nicht vor.

Ebenso wenig wurde untersucht, ob und inwiefern Inklusion und Bildungsstandards die tatsächliche Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung beeinflussen, wenngleich dies begründet angenommen werden kann. Denn die Ergebnisse früherer Studien weisen darauf hin, dass einerseits Situationen und Kontexte, die unterschiedliche Ziele der Leistungsbeurteilung akzentuieren (Holmeier, 2012; Wilbert & Gerdes, 2009) und andererseits auch Personenmerkmale der Lehrkraft (Mischo & Rheinberg, 1995; Retelsdorf & Günther, 2011) mit der BNO zusammenhängen bzw. diese beeinflussen. Da aus der sozialpsychologischen Forschung bekannt ist, dass sowohl Merkmale der Situation als auch Merkmale der Person Einfluss auf gezeigtes Verhalten nehmen können (Richard et al., 2003), erscheint es lohnend, auch Beurteilungsverhalten als Ergebnis einer Person-Situation-Interaktion zu betrachten. Übertragen auf die Anwendung von Bezugsnormen kann angenommen werden, dass eine situationelle Auseinandersetzung mit Inklusion oder Bildungsstandards je nach Einstellungen zu diesen Reformen unterschiedliches Beurteilungsverhalten hervorrufen kann. Diesbezüglich zeigt die sozialpsychologische Forschung, dass Einstellungen durch die Konfrontation mit den Einstellungsobjekten zugänglich gemacht und somit handlungswirksam werden können (Fazio et al., 1982; Fazio et al., 1986). Somit ist denkbar, dass aufgrund der

Auseinandersetzung mit Inklusion und Bildungsstandards jeweils vorhandene positive oder negative Einstellungen der Lehrkräfte zu diesen Reformen handlungswirksam werden und entweder ein mit Inklusion bzw. Bildungsstandards kongruentes oder ein dazu entgegengesetztes Beurteilungsverhalten im Sinne der Anwendung unterschiedlicher Bezugsnormen hervorrufen.

Schließlich wurde verdeutlicht, dass beide Reformen bei Lehrkräften das Einnehmen jeweils unterschiedlicher Perspektiven bei der Leistungsbeurteilung fördern können, da sie unterschiedliche Maßstäbe der Leistungsbeurteilung akzentuieren. Während mit Inklusion vordergründig ein individuell subjektiver Blick auf Schülerleistungen fokussiert wird (z.B. Seifert & Müller-Zastrau, 2014), lenken die Bildungsstandards und VERA das Hauptaugenmerk auf objektive, kriteriale Beurteilungsmaßstäbe. Ausgehend von Befunden, wonach bei Lehrkräften Stereotype über bestimmte Schülergruppen vorzuliegen scheinen (z.B. Glock & Krolak-Schwerdt, 2013; Nürnberger et al., 2016), wurde auf Basis der sozialpsychologischen Shifting-Standards-Theorie (Biernat et al., 1991; Biernat, 2009) verdeutlicht, dass Stereotype in der Leistungsbeurteilung auf objektiven Skalen stärker sichtbar werden, während sie in der Leistungsbeurteilung auf subjektiven Skalen eher verdeckt werden. Diese Annahmen der Shifting-Standards-Theorie konnten empirisch bereits vielfach gestützt werden (Biernat & Manis, 1994; Biernat & Fuegen, 2001; Biernat, 2009; Biernat et al., 2010) und scheinen im Kontext der Hervorhebung unterschiedlicher Beurteilungsmaßstäbe durch Inklusion und Bildungsstandards äußerst relevant. Denn das Denken in objektiven Beurteilungsmaßstäben kann demnach dazu führen, dass stereotypisierte Schülerinnen und Schüler (unbeabsichtigter Weise) unzutreffend beurteilt werden, wobei auch das mit Inklusion verbundene Denken in individuell subjektiven Maßstäben stereotypengeleitete Beurteilungen nicht verringern, sondern nur verdecken würde. Jedoch fehlen bisher Er-

kenntnisse dazu, inwiefern sich die jeweils durch Inklusion und Bildungsstandards hervor-  
gehobenen Beurteilungsmaßstäbe auf die Einschätzung von stereotypisierten gegenüber  
nicht stereotypisierten Schülergruppen auswirken können.

### **3.2 Fragestellungen und Studienüberblick**

Aus dem dargestellten theoretischen und empirischen Erkenntnisstand und den darin aufge-  
zeigten Lücken lassen sich folgende drei Fragestellungen ableiten:

1. Verbinden Lehrkräfte mit Inklusion und Bildungsstandards unterschiedliche  
Implikationen für Unterrichtsgestaltung und Leistungsbeurteilung?
2. Führt die Auseinandersetzung mit dem Thema Bildungsstandards oder mit dem  
Thema Inklusion bei Lehrkräften je nach Einstellung zu diesen Reformen zur  
Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung?
3. Welche Konsequenzen gehen mit der Anwendung subjektiver und objektiver  
Beurteilungsmaßstäbe für die Leistungsbeurteilung stereotypisierter Schüler-  
gruppen einher?

Zur Untersuchung dieser Fragestellungen wurden unter Rückgriff auf die erläuterten erzie-  
hungswissenschaftlichen und sozialpsychologischen Theorien drei empirische Studien kon-  
zipiert und durchgeführt.

Die erste Studie („*Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven  
und standardorientierten Unterricht aus der Sicht von Lehrkräften*“) geht der Frage nach,  
ob und inwiefern mit Inklusion und Bildungsstandards aus der Sicht von Lehrkräften eine  
jeweils andere Unterrichtsgestaltung und Leistungsbeurteilung verbunden ist. Dabei wird  
bzgl. der Unterrichtsgestaltung die theoretische Unterscheidung von Sicht- und  
Tiefenstrukturen des Unterrichts (Kunter & Trautwein, 2013; Oser & Baeriswyl, 2001) und  
bzgl. der Leistungsbeurteilung die Theorie der BNO (Rheinberg, 1980) aufgegriffen. Den  
Ausgangspunkt bilden einerseits die vielschichtig diskutierten, unterschiedlichen

Implikationen beider Reformen für Unterrichtsgestaltung und Leistungsbeurteilung (z.B. Seitz, 2006b; Teumer, 2012; Werning & Arndt, 2015) und andererseits erste empirische Hinweise darauf, dass Standardisierungsprozesse aus Sicht von Lehrkräften eine differenzierende Unterrichtsgestaltung einschränken können (Beer, 2006; Racherbäumer & Kühn, 2013).

Innerhalb einer experimentellen Studie wurden 135 Lehrkräfte im Vorbereitungsdienst dazu befragt, wie Merkmale auf Ebene der Sicht- und Tiefenstrukturen von Unterricht sowie die Leistungsbeurteilung in Form der drei BNO aus ihrer Sicht jeweils im inklusiven und im standardorientierten Unterricht ausgestaltet werden würden. Dabei wurden ebenfalls die Einstellungen zu und die Erfahrungen mit Inklusion und Bildungsstandards sowie Selbstwirksamkeitserwartungen bzgl. der Gestaltung inklusiven und standardorientierten Unterrichts berücksichtigt, da diese bei der Umsetzung beider Reformen eine wichtige Rolle zu spielen scheinen (Gebhardt et al., 2015; Grillitsch, 2010; Richter et al., 2014; Ruberg & Porsch, 2017). Da jedoch keine Erkenntnisse dazu vorliegen, auf welche Weise diese Lehrermerkmale konkret mit der Einschätzung von Unterrichtsgestaltung und Leistungsbeurteilung zusammenhängen, wurden diese als Kontrollvariablen behandelt. Weil Studie 1 erstmals überprüfte, inwiefern sich in der Literatur diskutierte unterschiedliche bzw. gegensätzliche Implikationen beider Reformen empirisch bestätigen lassen, wurde ein exploratives Vorgehen gewählt und auf die Formulierung gerichteter Hypothesen verzichtet. Ausgehend vom dargestellten bildungswissenschaftlichen Diskurs sollten sich jedoch Unterschiede zwischen der jeweils mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung zeigen, die auch unter Kontrolle der jeweils zugehörigen Einstellungen, Erfahrungen und Selbstwirksamkeitserwartungen bestehen bleiben.

Um zusätzlich bekannten Differenzen in der Beschreibung von Unterrichtsgestaltung, die sich zwischen unterschiedlichen Betrachtungsperspektiven von Unterricht zeigen (z.B. Clausen, 2002), Rechnung zu tragen, wurden in Studie 1 zwei Perspektiven auf Unterricht kontrastiert: die Einschätzung des eigenen Unterrichts und die Einschätzung des Unterrichts einer fremden (fiktiven) Lehrkraft. Beschreibungen des eigenen Unterrichts können im Sinne eines *self-presentation bias* (Kopcha & Sullivan, 2007) positiv verzerrt oder durch soziale Erwünschtheit beeinflusst sein (Wubbels, Brekelmans & Hooymayers, 1992). Eine solche allgemeine Tendenz zur Beschreibung „guten Unterrichts“ könnte Unterschiede, die sich aus Lehrersicht zwischen beiden Reformen zeigen, überdecken. Entsprechend sollten Unterschiede zwischen der jeweils mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung bei der Einschätzung des Unterrichts einer fremden Lehrkraft größer ausfallen als bei der Einschätzung des eigenen Unterrichts. Somit weist Studie 1 ein 2 x 2-faktorielles Design mit dem Innersubjektfaktor Unterrichtskontext (Inklusion vs. Bildungsstandards) und dem Zwischensubjektfaktor Beurteilungsperspektive (Selbsteinschätzung vs. Fremdeinschätzung) auf.

Da Inklusion vorrangig die individuelle Bezugsnorm bei der Leistungsbeurteilung impliziert, während mit Bildungsstandards eher die kriteriale und soziale Bezugsnorm hervorgehoben werden, untersucht die zweite Studie (*„Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung“*) die Frage, ob mit Inklusion und Bildungsstandards je nach Einstellungen zu diesen Reformen die Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung einhergeht. Den Ausgangspunkt bildet die Annahme, dass aufgrund der Auseinandersetzung mit Inklusion oder Bildungsstandards jeweils positive oder negative Einstellungen der Lehrkräfte zu diesen Reformen handlungswirksam werden und entweder ein mit Inklusion

bzw. Bildungsstandards kongruentes oder ein dazu entgegengesetztes Beurteilungsverhalten hervorgerufen wird. Davon ausgehend, wurde in einer Stichprobe von 57 Lehrkräften im Vorbereitungsdienst experimentell variiert, ob diese sich vor der Beurteilung von Schülerleistungen entweder mit dem Reformansatz „Bildungsstandards“ oder mit „Inklusion“ auseinandersetzten (situationeller Rahmen). Nach der Beurteilung der Schülerleistungen wurden die Einstellungen zu Inklusion und Bildungsstandards (als Personenmerkmale) erfragt. Somit wurden in Studie 2 Merkmale der Situation und der Person gemeinsam als Einflussfaktoren auf die BNO im Kontext von Inklusion und Bildungsstandards betrachtet. Bei Personen mit einer positiven Einstellung zu Bildungsstandards sollte eine Auseinandersetzung mit dieser Reform zu einer geringeren Ausprägung der individuellen BNO und zu einer stärkeren Ausprägung der kriterialen und sozialen BNO führen. Werden Bildungsstandards eher abgelehnt, sollte durch die Auseinandersetzung damit diese negative Einstellung handlungsleitend werden und entsprechend eine stärkere individuelle BNO und eine geringere kriteriale und soziale BNO bei der Leistungsbeurteilung auftreten. Ein spiegelbildliches Muster sollte sich bei der Auseinandersetzung mit Inklusion in Abhängigkeit von der diesbezüglichen Einstellung zeigen.

Die dritte Studie („*Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective*“) ist in zwei Teilstudien gegliedert und untersucht mögliche Konsequenzen der Anwendung subjektiver und objektiver Beurteilungsmaßstäbe für die Leistungsbeurteilung stereotypisierter Schülergruppen. Während Inklusion primär einen individuell subjektiven Blick auf Schülerleistungen impliziert, stellen Bildungsstandards objektive Beurteilungsmaßstäbe in den Vordergrund. Aus diesem Grund können beide Reformen bei Lehrkräften das Einnehmen jeweils unterschiedlicher Perspektiven bei der Leistungsbeurteilung fördern. Wie die Shifting-

Standards-Theorie (Biernat et al., 1991; Biernat, 2009) verdeutlicht, kann sich das Einnehmen einer subjektiven oder objektiven Beurteilungsperspektive unterschiedlich auf die Beurteilung stereotypisierter Schülergruppen auswirken (z.B. Biernat & Manis, 1994). Davon ausgehend und auf Basis von Befunden, dass bei Lehrkräften Stereotype über bestimmte Schülergruppen vorzuliegen scheinen (z.B. Glock & Krolak-Schwerdt, 2013; Nürnberger et al., 2016) untersucht die erste Teilstudie ( $N = 115$  Lehramtsstudierende) das Geschlechterstereotyp, dass Jungen in Mathematik talentierter sind als Mädchen (Li, 1999). In der zweiten Teilstudie ( $N = 265$  Lehramtsstudierende) wird das ethnische Stereotyp untersucht, dass türkischstämmige Lernende schlechtere Schulleistungen erbringen, als jene ohne Migrationshintergrund (Lorenz et al., 2016).

In zwei  $2 \times 2$ -faktoriellen, experimentellen Vignettenstudien wurde jeweils überprüft, ob der Einsatz einer subjektiven Beurteilungsskala (im Kontext eines individualisierten Beurteilungssettings) und der Einsatz einer objektiven Beurteilungsskala (im Kontext eines standardisierten Beurteilungssettings) jeweils zu unterschiedlichen Beurteilungen von stereotypisierten Lernenden gegenüber nicht stereotypisierten Lernenden führt. Anhand fiktiver, identischer Leistungsbeschreibungen waren in der ersten Teilstudie Mathematikleistungen eines Mädchens oder eines Jungens einzuschätzen und in der zweiten Teilstudie Deutschleistungen eines Schülers mit türkischem oder deutschem Namen. Im individualisierten Beurteilungssetting wurde die fiktive Leistungsbeschreibung als Ergebnis eines individuellen Portfolios dargestellt, welches verschiedene vorausgegangene Schülerleistungen dokumentiert. Die Beurteilung erfolgte auf einer subjektiven Skala von „sehr schlecht“ bis „sehr gut“. Im standardisierten Beurteilungssetting wurde die fiktive Leistungsbeschreibung als Ergebnis der VERA dargestellt. Die Beurteilung erfolgte anhand der Messskala der Bildungsstandards. Sowohl die Beurteilungsskala (subjektiv vs. objektiv) als auch die zu beurteilende Person (Mädchen

vs. Junge; türkischer Schüler vs. deutscher Schüler) wurden dabei als Zwischensubjektfaktoren variiert.

Gemäß der Shifting-Standards-Theorie (Biernat et al., 1991; Biernat & Manis, 1994) sollten Mitglieder der jeweils stereotypisierten Schülergruppe (Mädchen in Mathematik und türkischstämmige Schüler in Deutsch) gegenüber Mitgliedern der jeweils nicht stereotypisierten Schülergruppe (Jungen in Mathematik und deutschstämmige Schüler in Deutsch) negativer beurteilt werden, wenn eine objektive Beurteilungsskala eingesetzt wird. Bei Verwendung einer subjektiven Skala, deren Ausprägungen unterschiedlich interpretiert werden können, sollten sich hingegen keine Unterschiede in der Leistungsbeurteilung beider Schülergruppen zeigen.



## 4 Studien

### 4.1 Studie 1: Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven und standardorientierten Unterricht aus der Sicht von Lehrkräften

Anmerkung: Dieser Teil der Dissertation entspricht einem Manuskript, welches zur Publikation eingereicht wurde.

Holder, K. & Kessels, U. (2018). Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven und standardorientierten Unterricht aus der Sicht von Lehrkräften. *Zur Publikation eingereichtes Manuskript.*

**Die finale Publikation ist verfügbar bei Springer unter/**

**The final publication is available at [link.springer.com](https://link.springer.com) via**

**<https://doi.org/10.1007/s11618-018-0854-7>**

### **Zusammenfassung**

Die gleichzeitige Umsetzung von Inklusion und Bildungsstandards geht mit einem erziehungswissenschaftlichen Diskurs um deren Vereinbarkeit einher. Jedoch liegt kaum Evidenz dazu vor, ob und inwiefern Lehrkräfte mit diesen Reformen gegensätzliche Anforderungen verbinden. Die vorliegende experimentelle Studie untersucht, ob und inwiefern Lehrkräfte im Vorbereitungsdienst ( $N = 135$ ) mit Inklusion und Standardorientierung eine jeweils andere Unterrichtsgestaltung und Leistungsbeurteilung verbinden, wobei zwei Perspektiven auf Unterricht (Selbst- und Fremdeinschätzung) kontrastiert werden. Die Ergebnisse zeigten, dass sich aus Lehrersicht inklusiver Unterricht von standardorientiertem Unterricht hinsichtlich der Umsetzung zentraler Merkmale von Unterrichtsqualität und hinsichtlich der Bezugsnormen bei der Leistungsbeurteilung unterscheidet. Diese Unterschiede sind zum Teil auf Einstellungen zu und Erfahrungen mit beiden Reformen sowie Selbstwirksamkeitserwartungen zurückzuführen. Der Unterricht in beiden Kontexten wird deutlich unterschiedlicher beschrieben, wenn nicht das eigene, sondern das Unterrichtshandeln einer anderen (fiktiven) Lehrkraft charakterisiert wird.

Schlüsselwörter: Inklusion, Bildungsstandards, Unterrichtsgestaltung, Leistungsbeurteilung

### **Abstract**

In educational systems in German-speaking countries, at present, two main reforms are simultaneously being implemented: inclusion and educational standards. While it is discussed in the literature that these two reforms may have contradictory implications for teaching, only few empirical studies have actually addressed this issue from a teachers' perspective. The present study uses an experimental design to investigate how pre-service teachers ( $N = 135$ ) perceive instructional quality and performance evaluation in the context of inclusion and educational standards respectively, taking account two perspectives on teaching: one's own teaching versus other's teaching. Results show that pre-service teachers associate different forms of instructional quality and different reference norm orientations with inclusion in contrast to educational standards. Attitudes towards both reforms, experiences with them, and self-efficacy beliefs partly explained these differences. Further, inclusive instruction and standards-based instruction are perceived as differing more when the teaching of a (fictitious) other teacher is described compared to describing one's own teaching.

**Keywords:** educational standards, inclusion, instructional quality, performance evaluation

## **1 Einleitung**

Mit der Implementation der Bildungsstandards einerseits und der zunehmenden Realisierung von Inklusion andererseits treffen in deutschsprachigen Schulsystemen zwei Reformen aufeinander, die den aktuellen bildungswissenschaftlichen Diskurs wesentlich bestimmen (Eversheim 2015). Dabei wird angemerkt, dass mit der gleichzeitigen Implementation beider Reformansätze paradoxe Anforderungen an Lehrkräfte einhergehen können (Teumer 2012), die das damit verbundene Spannungsverhältnis zwischen individueller Förderung und der Vorgabe einheitlicher Lernziele ausbalancieren müssten.

Es werden sowohl unterschiedliche Implikationen beider Reformen für Unterrichtsgestaltung (z.B. Seitz und Platte 2006) als auch für Leistungsbeurteilung (z.B. Seitz 2006) diskutiert. Jedoch liegt kaum Evidenz vor, inwiefern sich Inklusion und Standardorientierung in der Wahrnehmung von Lehrkräften unterschiedlich auf ihre praktische Tätigkeit auswirken. Aus diesem Grund geht die vorliegende experimentelle Studie der Frage nach, ob und inwiefern mit Inklusion und Standardorientierung aus Sicht von Lehrkräften im Vorbereitungsdienst eine jeweils andere Unterrichtsgestaltung und Leistungsbeurteilung verbunden ist. Dabei werden zwei Perspektiven auf Unterrichtshandeln eingenommen (Selbst- und Fremdeinschätzung), um bekannten Differenzen zwischen verschiedenen Urteilspektiven (Clausen 2002) Rechnung zu tragen.

## **2 Die (Un-)Vereinbarkeit von Inklusion und Standardorientierung**

Inwiefern die gleichzeitige Implementation von Inklusion und Bildungsstandards gegensätzliche Anforderungen impliziert, wird kontrovers diskutiert (z.B. Biewer 2012; Teumer 2012). Einerseits wird argumentiert, dass Bildungsstandards den im Kontext von Inklusion besonders betonten Individualisierungsprozessen nicht entgegenstehen, sondern diese sogar unterstützen: Da lediglich die Lernziele, nicht aber die Lernprozesse standardisiert seien, entstünden Freiräume für individualisiertes Lernen (z.B. KMK 2010).

Seitens der Sonder- und Inklusionspädagogik wird hingegen kritisch angemerkt, dass es im Kontext der Standardorientierung vorrangig um das Erfüllen einer Leistungsnorm statt um eine bestmögliche Entwicklung der Lernenden gehe (Eversheim, 2015).

Erste Studienergebnisse aus England zeigten, dass Lehrkräfte Spannungen zwischen Inklusion und der Standard-Agenda wahrnehmen (z.B. Ainscow et al. 2006b). Auch wenn im deutschsprachigen Raum diskutiert wird, dass durch Bildungsstandards „die Etablierung eines inklusiven Bildungssystems beeinträchtigt“ (Biewer 2012, S.16) würde, sind entsprechende Forschungsergebnisse bisher rar.

### **2.1 Unterrichtsgestaltung im Kontext von Inklusion und Standardorientierung**

Die Implementation von Inklusion und Bildungsstandards zielt jeweils darauf ab, Unterricht zu verändern (KMK 2010, 2011). Allerdings ergeben sich aus den Bildungsstandards selbst keine direkten Handlungshinweise (Wacker et al. 2013), sondern erst aus ergänzenden Handreichungen zur Unterrichtsgestaltung (vgl. Drieschner 2009; KMK 2010). Ähnlich verhält es sich mit Inklusion, wobei erste Leitlinien inklusiver Didaktik beschrieben werden (z.B. Kullmann et al. 2014).

Inklusiver Unterricht soll die „Leitidee individueller Herausforderung in sozialer Eingebundenheit“ (Seitz und Scheidt 2012, o. S.) verwirklichen. Individueller Förderung und Differenzierung kommen folglich besondere Bedeutung zu, wobei gleichzeitig der Fokus auf die Herstellung von Gemeinsamkeit gelegt werden soll (Kullmann et al. 2014). Studien zur Umsetzung inklusiven Unterrichts zeigten, dass dieser sich nicht grundsätzlich von alltäglich verbreitetem Unterricht unterscheidet (z.B. Davis und Florian 2004), sich jedoch Merkmale wie eine effiziente Klassenführung (Emmer und Stough 2001) und die Unterstützung bei der Anwendung von Lernstrategien (Swanson 1999) als besonders bedeutsam für die Förderung von Kindern mit Lernschwierigkeiten herausstellten (vgl. zsf. Werning

2014). Für gelingende inklusive Lehr-Lernprozesse werden zudem Selbstwirksamkeitserwartungen, Einstellungen und Erfahrungen von Lehrkräften als bedeutsam angesehen (z.B. Gebhardt et al. 2015).

Mit der Implementation der Bildungsstandards ist das Ziel verbunden, Unterricht kompetenzorientiert auszurichten (z.B. Lenski et al. 2015). Damit soll ein hohes Maß an kognitiver Aktivierung und Handlungsorientierung (Drieschner 2009; Pöhlmann et al. 2014) sowie eine Differenzierung nach individuellen Voraussetzungen (z.B. KMK 2010) einhergehen. Forschungsergebnisse aus Deutschland zeigten, dass mit der Implementation der Bildungsstandards vereinzelt Veränderungen des Unterrichts einhergehen, jedoch keine substantiellen Veränderungen stattfinden (z.B. Wacker et al. 2013). Ähnlich wie bei Inklusion gelten auch die Einstellungen zu und die Erfahrungen mit Bildungsstandards als zentral für eine gelingende Implementation (Grillitsch 2010). Bezogen auf die standardbasierten Vergleichsarbeiten (VERA) zeigten z.B. Richter et al. (2014), dass Lehrkräfte, die VERA als Instrument zur Unterrichtsentwicklung ansehen, ihren Unterricht kompetenzorientierter und differenzierter gestalten.

Zusammenfassend werden in der Literatur zwar jeweils unterschiedliche Aspekte der Unterrichtsgestaltung hervorgehoben, letztlich wäre jedoch in beiden Kontexten „guter Unterricht“ im Sinne der Basisdimensionen von Unterrichtsqualität (z.B. Kunter et al. 2011) anzustreben. Erste Studien deuten jedoch darauf hin, dass sich die beiden Reformansätze diesbezüglich konträr gegenüberstehen. Lehrkräfte berichteten, dass sie die Umsetzung von Differenzierung im Unterricht im Kontext standardisierter Abschlussprüfungen als eingeschränkt wahrnehmen (Racherbäumer und Kühn 2013) und dass sie mehrheitlich keine verbesserten Möglichkeiten der Individualisierung im Unterricht durch die Standards sehen (Beer 2006).

Eine offene Frage ist, ob und inwiefern Lehrkräfte mit Inklusion und Standardorientierung eine unterschiedliche Unterrichtsgestaltung verbinden. Der Diskurs um ein Spannungsverhältnis beider Reformansätze lässt dieses vermuten (Biewer 2012), ohne dass hierzu ausreichende Evidenz vorliegt.

## **2.2 Leistungsbeurteilung im Kontext von Inklusion und Standardorientierung**

Als ein zentraler Unterschied zwischen beiden Reformen werden die jeweiligen Perspektiven auf Leistung thematisiert (Ainscow et al. 2006a). Dabei wird betont, dass Inklusion vordergründig den Lernprozess, Bildungsstandards dagegen Lernergebnisse fokussieren (Biewer 2012). Die mit den Standards verbundenen punktuellen Leistungsmessungen stünden zudem prozessorientierten Formen der Leistungsbeurteilung gegenüber (Seitz 2006).

Zudem werden Inklusion und Bildungsstandards zu unterschiedlichen Bezugsnormen bei der Leistungsbeurteilung in Beziehung gesetzt (Halbheer und Reusser 2008; Wilhelm 2012). Nach Rheinberg (1980) werden die individuelle, die soziale und die kriteriale Bezugsnorm unterschieden und die tendenzielle Verwendung einer Bezugsnorm als Bezugsnormorientierung (BNO) der Lehrkraft definiert. Im Kontext von Inklusion wird eine primäre Orientierung an der individuellen Bezugsnorm betont, da „jeder Entwicklungsfortschritt [...] als Leistung bewertet“ (Wilhelm 2009, S. 107) werden soll. Bildungsstandards implizieren hingegen primär die kriteriale Bezugsnorm, da sie externe Kriterien zur Lernzielüberprüfung bereitstellen (KMK 2010) sowie die soziale Bezugsnorm, da die Ergebnisrückmeldungen zu VERA üblicherweise die Leistungen eines Kindes zu einer sozialen Vergleichsgruppe klassen- oder schulübergreifend (Helmke und Hosenfeld 2003) und bspw. in Berlin auch klassenintern (z.B. Emmrich et al. 2012) in Relation setzen.

Für Lehrkräfte ergibt sich insgesamt die Herausforderung, sowohl individuelle Fortschritte anzuerkennen als auch übergreifende Kriterien in die Leistungsbeurteilung einzu-

beziehen (Bräu 2008). Empirische Befunde zur Wahrnehmung unterschiedlicher Anforderungen an die Leistungsbeurteilung durch Inklusion und Standardorientierung durch Lehrkräfte sind jedoch rar. Eine erste experimentelle Studie zur tatsächlichen Anwendung von Bezugsnormen zeigte, dass die Beschäftigung mit dem Konzept der Bildungsstandards bei Lehrkräften mit einer positiven Einstellung zu diesen eine geringere individuelle BNO und eine stärkere soziale BNO hervorrief (Holder und Kessels im Druck).

### **3 Selbst- und Fremdeinschätzung von Unterrichtsgestaltung und Leistungsbeurteilung**

Unterrichtsgestaltung und Leistungsbeurteilung (BNO) können durch Selbst- oder Fremdeinschätzung beschrieben werden, wobei in der Unterrichtsforschung am häufigsten Selbstberichte von Lehrkräften oder Fremdeinschätzungen durch externe Beobachter oder Schüler/-innen herangezogen werden (z.B. Clausen 2002). Hierbei fallen die Übereinstimmungen zwischen Selbstberichten von Lehrkräften und anderen Urteilsperspektiven insgesamt gering bis moderat aus (Clausen 2002; Lenski et al. 2015).

Als mögliche Erklärung der Übereinstimmungsdifferenzen wird u.a. angeführt, dass Lehrerurteile über eigenen Unterricht durch selbstwertdienliche oder sozial erwünschte Einschätzungen verzerrt sein können (Wubbels et al. 1992; Clausen 2002). Ein aus der Sozialpsychologie bekanntes, ähnliches Phänomen ist der sog. *self-presentation bias* (Kopcha und Sullivan 2007). Dieser bezeichnet die Tendenz von Menschen, ihr Handeln und sich selbst nach außen möglichst positiv zu beschreiben. So berichteten Lehrkräfte bspw. häufiger, im Unterricht erwünschte Methoden anzuwenden, als ihre Schüler/-innen (Bober et al. 1998).

Auch bzgl. Leistungsbeurteilung konnten Fischbach et al. (2015) zeigen, dass die Einschätzung der BNO von Lehrkräften durch Schüler/-innen und durch sie selbst nur moderat übereinstimmt. Da bestimmte BNO hinsichtlich ihrer Auswirkungen auf



Schülermerkmale pädagogisch erwünschter sind als andere (zu positiven Auswirkungen der individuellen BNO vgl. Köller 2005) ist ein *self-presentation bias* bei Selbstberichten zur BNO ebenfalls denkbar.

Aussagen über eigenen Unterricht können im Sinne eines *self-presentation bias* positiv verzerrt oder durch soziale Erwünschtheit beeinflusst sein, was als allgemeine Tendenz etwaige Unterschiede im mit Inklusion und Standortorientierung verbundenen Unterricht überdecken könnte. Entsprechend sollten Einschätzungen über das eigene Unterrichten ergänzt werden durch Einschätzungen des Unterrichts anderer Personen, um mögliche Unterschiede aufzeigen zu können.

#### **4 Fragestellung und Hypothesen**

Ziel der vorliegenden experimentellen Studie ist es, die mit Inklusion und Bildungsstandards verbundene Unterrichtsgestaltung und Leistungsbeurteilung aus Sicht von Lehrkräften im Vorbereitungsdienst zu untersuchen, um zu prüfen, ob und inwiefern sich die in der Literatur diskutierte Gegensätzlichkeit (z.B. Seitz und Platte 2006) empirisch bestätigen lässt. Dies erfolgt anhand zentraler Merkmale von Unterrichtsqualität, die sowohl auf Tiefen- als auch auf Sichtstrukturen des Unterrichts (Oser und Baeriswyl 2001) verweisen, sowie anhand der drei Formen der BNO (Rheinberg 1980).

Dabei werden zwei unterschiedliche Perspektiven gewählt. Einerseits wird nach der eigenen unterrichtlichen Praxis in beiden Kontexten gefragt. Wegen der oben beschriebenen möglichen Verzerrungen der Sicht auf den eigenen Unterricht, wird andererseits erhoben, wie nach Meinung der teilnehmenden Lehrkräfte die Unterrichtsgestaltung und die Leistungsbeurteilung einer anderen (fiktiven) Lehrkraft aussehen würde, die entweder inklusiv oder standardorientiert unterrichtet. Hierdurch soll unabhängig vom *self-presentation bias* (Kopcha und Sullivan 2007) und von sozialer Erwünschtheit erfasst werden, welche Anforde-

rungen Lehrkräfte an Unterrichtsgestaltung und Leistungsbeurteilung mit den beiden Reformansätzen verbinden. Entsprechend vermuten wir, dass etwaige Unterschiede zwischen den beiden Unterrichtskontexten deutlicher hervortreten, wenn das jeweilige Unterrichten einer anderen (fiktiven) Lehrkraft beschrieben wird als wenn der eigene Unterricht eingeschätzt wird.

Als Kontrollvariablen werden Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie Selbstwirksamkeitserwartungen bzgl. der Gestaltung inklusiven und standardorientierten Unterrichts erhoben, da diese bei der Umsetzung beider Reformen eine zentrale Rolle zu spielen scheinen (Grillitsch 2010; Gebhardt et al. 2015).

Da mit unserer Studie erstmals die Frage untersucht wird, ob und inwiefern mit Inklusion und Bildungsstandards aus Sicht von Lehrkräften unterschiedliche Implikationen für Unterrichtshandeln einhergehen, werden folgende, zum Teil ungerichtete Hypothesen mit explorativem Charakter aufgestellt:

1. Aus Sicht von Lehrkräften unterscheidet sich
  - a. die Unterrichtsgestaltung im inklusiven Unterricht von jener im standardorientierten Unterricht hinsichtlich zentraler Merkmale von Unterrichtsqualität.
  - b. die Leistungsbeurteilung im inklusiven Unterricht von jener im standardorientierten Unterricht hinsichtlich der drei BNO.

2. Die Unterschiede zwischen der mit Inklusion und Bildungsstandards verbundenen  
a. Unterrichtsgestaltung hinsichtlich zentraler Merkmale von Unterrichtsqualität

und

b. Leistungsbeurteilung hinsichtlich der drei BNO

bleiben auch unter Kontrolle der Einstellungen zu Inklusion und zu Bildungsstandards, der Erfahrungen mit Inklusion und Bildungsstandards und der Selbstwirksamkeitserwartung bezogen auf inklusiven und standardorientierten Unterricht bestehen.

3. Die Unterschiede zwischen der mit Inklusion und Bildungsstandards verbundenen

a. Unterrichtsgestaltung hinsichtlich zentraler Merkmale von Unterrichtsqualität

und

b. Leistungsbeurteilung hinsichtlich der drei BNO

variieren je nach Beurteilungsperspektive: Sie fallen bei der Einschätzung des Unterrichts einer fremden Lehrkraft größer aus als bei der Einschätzung des eigenen Unterrichts.

## 5 Methode

Die Hypothesen wurden in einer experimentellen Studie mit einem 2 (*Unterrichtskontext*) x 2 (*Beurteilungsperspektive*) Design überprüft. Der Unterrichtskontext (Inklusion vs. Bildungsstandards) wurde dabei innerhalb der Versuchspersonen (Vpn) und die Beurteilungsperspektive (Selbsteinschätzung vs. Fremdeinschätzung) zwischen den Vpn variiert. Nachfolgend werden die Untersuchungsdurchführung, die Erhebungsinstrumente, die Stichprobe sowie die Datenanalyse beschrieben.

### 5.1 Untersuchungsdurchführung und Erhebungsinstrumente

Die Vpn nahmen im Rahmen eines schulpraktischen Seminars freiwillig an einer schriftlichen Befragung zum Thema „Unterrichten im Kontext aktueller Entwicklungen“ teil, in die folgende Erhebungsinstrumente eingebettet waren.

Als Erstes wurden allen Vpn zwei kurze Texte präsentiert, in denen Inklusion und Bildungsstandards als Rahmenbedingungen des Unterrichts überblicksartig beschrieben wurden. Diese waren auf Basis einschlägiger Publikationen entwickelt worden und enthielten allgemeine Informationen zur Einführung beider Reformansätze und damit verbundener Ziele. Bewusst wurden keine Angaben zu Unterrichtsgestaltung oder Leistungsbeurteilung im Konkreten gemacht.

Anschließend beschrieben die Vpn anhand zentraler Merkmale von Unterrichtsqualität, wie Unterricht im jeweiligen Kontext (*Inklusion vs. Bildungsstandards*) jeweils gestaltet werden würde und anhand der drei BNO, wie Leistungsbeurteilung jeweils vorgenommen werden würde. Im Einzelnen gaben sie an, inwiefern Merkmale auf Ebene der Sichtstrukturen (Methodenvielfalt, Frontalunterricht, Differenzierung) und auf Ebene der Tiefenstrukturen von Unterricht (kognitive Aktivierung, Unterstützung durch die Lehrkraft, effiziente Klassenführung) sowie verschiedene Bezugsnormen bei der Leistungsbeurteilung (individuelle, soziale und kriteriale) ihrer Meinung nach jeweils umgesetzt würden (vgl. Tab. 1).

Dabei wurde die Beurteilungsperspektive (*Selbsteinschätzung vs. Fremdeinschätzung*) als Zwischensubjektfaktor variiert. Die Vpn sollten entweder beschreiben, inwieweit sie selbst die beschriebenen Merkmale von Unterrichtsqualität, und die drei BNO im jeweiligen Unterrichtskontext ausgestalten („*Wie gestalten Sie Ihren Unterricht jeweils, um standardorientiert (bzw. inklusiv) zu unterrichten?*“) oder inwieweit eine fremde (fiktive) Lehr-

---

kraft diese jeweils umsetzen würde („*Wie würde eine Lehrkraft den Unterricht jeweils gestalten, um die beschriebenen Ideen standardorientierten (bzw. inklusiven) Unterrichts umzusetzen?*“).

Anschließend wurden die Einstellungen der Teilnehmenden zu und ihre Erfahrungen mit Inklusion und Bildungsstandards sowie ihre Selbstwirksamkeitserwartungen bzgl. inklusivem und standardorientiertem Unterricht erhoben (vgl. Tab. 1). Tabelle 1 gibt eine Übersicht über die eingesetzten Instrumente. Für alle Skalen wurde pro Vpn ein Mittelwert aus den zugehörigen Items gebildet, wobei negativ gepolte Items vorher umkodiert wurden, so dass höhere Werte stärkeren Ausprägungen der Variablen entsprechen. Die Reliabilitäten der einzelnen Skalen fallen überwiegend zufriedenstellend bis gut aus, nur die Skala Einstellungen zu Bildungsstandards weist eine eher geringe Reliabilität auf (Cronbachs  $\alpha = .56$ ).

Tabelle 1  
Deskriptiva der eingesetzten Instrumente

Instrument	Items	Beispielitem	Alpha	M	SD	Antwortformat	Angepasste Items aus (Quelle)
Methodenvielfalt	6	„Wie oft würden Sie die folgenden pädagogischen bzw. organisatorischen Konzepte verfolgen: Arbeit mit kleinen Schüler/innengruppen				1 (sehr selten) -5 (sehr häufig)	Skala „Methodenvielfalt“ (Wagner et al. 2009)
		... im inklusiven Unterricht.“	.81	3,23	0,87		
		... im an Bildungsstandards orientierten Unterricht.“	.77	3,24	0,76		
Frontalunterricht	1	„Wie oft würden Sie die folgenden pädagogischen bzw. organisatorischen Konzepte verfolgen: Frontalunterricht (fragend-entwickelnder Unterricht)				1 (sehr selten) -5 (sehr häufig)	Eigenkonstruktion
		... im inklusiven Unterricht.“	-	2,83	1,11		
		... im an Bildungsstandards orientierten Unterricht.“	-	3,44	0,91		
Differenzierung	4	„Ich gebe Schülern/innen je nach Leistungsstärke unterschiedlich schwierige Aufgaben*“				1 (sehr selten) -5 (sehr häufig)	Skala „Differenzierung“ (Wagner et al. 2009)
		... im inklusiven Unterricht.“	.84	3,94	0,90		
		... im an Bildungsstandards orientierten Unterricht.“	.72	3,65	0,75		

Tabelle 1 (Fortsetzung)

Instrument	Items	Beispielitem	Alpha	M	SD	Antwortformat	Angepasste Items aus (Quelle)
Kognitive Aktivierung	4	„Ich stelle oft Fragen, die man nicht spontan beantworten kann, sondern die zum Nachdenken zwingen“ <sup>*</sup>				1 (sehr selten)	Skala „Kognitiv aktivierende Aufgaben bei der Einführung eines neuen Sachverhaltes und bei üben“ (2 Items; PISA Konsortium Deutschland 2006) u. Skala „Selbstständigkeit und Begründungspflicht beim Bearbeiten von Aufgaben/Kognitive Selbstständigkeit“ (2 Items; PISA Konsortium Deutschland 2006)
		... im inklusiven Unterricht.“	.84	3,08	0,88	-5 (sehr häufig)	
		... im an Bildungsstandards orientierten Unterricht.“	.78	3,50	0,72		
Unterstützung durch die Lehrkraft	4	„Ich unterstütze die Schüler/innen zusätzlich, wenn sie Hilfe brauchen“ <sup>*</sup>				1 (sehr selten)	Skala „Unterstützung durch den Lehrer im Mathematikunterricht“ (PISA Konsortium Deutschland 2006)
		... im inklusiven Unterricht.“	.88	4,44	0,74	-5 (sehr häufig)	
		... im an Bildungsstandards orientierten Unterricht.“	.74	4,33	0,56		
Effiziente Klassenführung	4	„Im Unterricht ist es laut und alles geht durcheinander“				1 (sehr selten)	Skala „Klassenführung-Disziplinprobleme im Mathematikunterricht“ (PISA Konsortium Deutschland 2006)
		... im inklusiven Unterricht.“	.90	3,05	0,89	-5 (sehr häufig)	
		... im an Bildungsstandards orientierten Unterricht.“	.88	3,22	0,76		

87

Tabelle 1 (Fortsetzung)

Instrument	Items	Beispielitem	Alpha	M	SD	Antwortformat	Angepasste Items aus (Quelle)
individuelle BNO	2	„Ich meine mit „schlechter Leistung“ ein Ergebnis, das unter dem vorherigen Resultat des jeweiligen Kindes liegt*“				1 (trifft überhaupt nicht zu)	Skala „Leistungsvergleich“ (FEBO, Rheinberg 1980)
		... im inklusiven Unterricht.“	.74	2,92	0,78	- 4 (trifft voll und ganz zu)	
		... im an Bildungsstandards orientierten Unterricht.“	.72	2,70	0,73		
soziale BNO	2	„Ich meine mit „schlechter Leistung“ ein Ergebnis, das deutlich unter dem Klassendurchschnitt liegt*“				1 (trifft überhaupt nicht zu)	Skala „Leistungsvergleich“ (FEBO, Rheinberg 1980)
		... im inklusiven Unterricht.“	.58	2,23	0,73	- 4 (trifft voll und ganz zu)	
		... im an Bildungsstandards orientierten Unterricht.“	.74	2,71	0,78		
kriteriale BNO	2	„Ich meine mit „schlechter Leistung“ ein Ergebnis, das unter den gestellten Anforderungen liegt*“				1 (trifft überhaupt nicht zu)	Eigenkonstruktion
		... im inklusiven Unterricht.“	.61	2,80	0,79	- 4 (trifft voll und ganz zu)	
		... im an Bildungsstandards orientierten Unterricht.“	.67	3,24	0,65		
Einstellungen zu Inklusion	4	„Ein gemeinsamer Unterricht behinderter und nicht behinderter Schüler/innen kann durch entsprechende Methoden allen Schülern/innen gerecht werden.“	.73	2,67	0,62	1 (trifft überhaupt nicht zu) - 4 (trifft voll und ganz zu)	Skala „Einstellung zur Gestaltung inklusiven Unterrichts“ (Bosse und Spörer 2014)



Tabelle 1 (Fortsetzung)

Instrument	Items	Beispielitem	Alpha	M	SD	Antwortformat	Angepasste Items aus (Quelle)
Einstellungen zu Bildungsstandards	6	„ <i>Bildungsstandards geben Klarheit, worauf es in der Schule letztendlich ankommt.</i> “	.56	2,35	0,43	1 (trifft überhaupt nicht zu) - 4 (trifft voll und ganz zu)	Skala „Einstellungen zu Bildungsstandards im Allgemeinen“ (Freudenthaler und Specht 2005)
Erfahrung mit Inklusion	2	„ <i>Wie viel haben/hatten Sie in Ihrer praktischen Tätigkeit mit Inklusion zu tun?</i> “	.84	3,07	1,09	1 (gar nicht) - 5 (sehr viel)	Eigenkonstruktion
Erfahrung mit Bildungsstandards	2	„ <i>Wie viel haben/hatten Sie in Ihrer praktischen Tätigkeit mit Bildungsstandards und/oder Vergleichsarbeiten (VERA) zu tun?</i> “	.78	2,72	0,98	1 (gar nicht) - 5 (sehr viel)	Eigenkonstruktion
Selbstwirksamkeitserwartung bezogen auf inklusiven Unterricht	3	„ <i>Ich traue mir zu, Unterricht so zu organisieren, dass auch Schüler/-innen mit Förderbedarf in ihrem eigenen Lerntempo zum Ziel kommen können.</i> “	.86	2,50	0,67	1 (trifft überhaupt nicht zu) - 4 (trifft voll und ganz zu)	Skala „Selbstwirksamkeit bezogen auf die Gestaltung inklusiven Unterrichts“ (Bosse und Spörer 2014)
Selbstwirksamkeitserwartung bezogen auf standardorientierten Unterricht	3	„ <i>Ich weiß, dass ich ein Unterrichtsthema so aufbereiten kann, dass die Schüler/-innen die in den Bildungsstandards beschriebenen Kompetenzen aufbauen können.</i> “	.75	3,10	0,41	1 (trifft überhaupt nicht zu) - 4 (trifft voll und ganz zu)	Eigenkonstruktion

\* Angegeben ist hier die Formulierung des Items in der Bedingung *Selbsteinschätzung* „Ich ...“. In der Bedingung *Fremdeinschätzung* wurde die jeweilige Formulierung entsprechend geändert zu „Die Lehrkraft...“.

## 5.2 Stichprobe

Die Stichprobe bestand aus  $N= 154$  Lehrkräften im Vorbereitungsdienst. 19 Vpn wurden aus den Analysen ausgeschlossen, da sie entweder lediglich Angaben zu einem der beiden Unterrichtskontexte gemacht hatten oder die Angaben aufgrund von Kommentierungen oder unklaren Kreuzsetzungen nicht auswertbar waren. Somit bestand die Analysestichprobe aus  $N= 135$  Lehrkräften im Vorbereitungsdienst (91 weiblich, 1 fehlende Angabe), wovon  $n= 58$  randomisiert der Bedingung Selbsteinschätzung und  $n= 77$  der Bedingung Fremdeinschätzung zugewiesen worden waren. Das durchschnittliche Alter der Vpn betrug  $M= 31,87$  ( $SD= 7,97$ ) Jahre und sie befanden sich im Durchschnitt seit ca. zehn Monaten im Vorbereitungsdienst ( $M= 9,75$ ;  $SD= 3,48$ ). Die Vpn wurden nach Studienende über deren Ziele und Ergebnisse aufgeklärt.

## 5.3 Datenanalyse

Die Hypothesen 1a. und 1b. wurden jeweils mit einer einfaktoriellen, multivariaten Varianzanalyse (MANOVA) überprüft, in die der *Unterrichtskontext* (Inklusion vs. Bildungsstandards) als Innersubjektfaktor einging. Die abhängigen Variablen bildeten einerseits (Hypothese 1a.) die Merkmale von Unterrichtsqualität (Methodenvielfalt, Frontalunterricht, Differenzierung, kognitive Aktivierung, Unterstützung durch die Lehrkraft, effiziente Klassenführung) und andererseits (Hypothese 1b.) die BNO (individuelle, soziale, kriteriale).

Anschließend wurden zur Überprüfung von Hypothese 2a. und 2b. zusätzlich die Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie die Selbstwirksamkeitserwartung bzgl. inklusivem und standardorientiertem Unterricht als Kovariaten in die Analysen eingeführt (MANCOVA).

Die Hypothesen 3a. und 3b. wurden jeweils mittels einer 2 (*Unterrichtskontext*: Inklusion vs. Bildungsstandards) x 2 (*Beurteilungsperspektive*: Selbsteinschätzung vs. Fremdeinschätzung) faktoriellen, multivariaten Kovarianzanalyse (MANCOVA) mit Einbezug der oben genannten Kovariaten geprüft, wobei der *Unterrichtskontext* als Innersubjektfaktor, die *Beurteilungsperspektive* als Zwischensubjektfaktor sowie deren Interaktion als unabhängige Variablen aufgenommen wurden. Die abhängigen Variablen bildeten ebenfalls einerseits die Merkmale von Unterrichtsqualität, (Hypothese 3a.) und andererseits die BNO (Hypothese 3b.).

## 6 Ergebnisse

### 6.1 Deskriptive Ergebnisse

Mittelwerte und Standardabweichungen der mit inklusivem und standardorientiertem Unterricht verbundenen Unterrichtsgestaltung bzgl. der verschiedenen Merkmale von Unterrichtsqualität, der BNO sowie der Kovariaten zeigt Tabelle 1.

Zunächst wurde mittels t-Tests überprüft, ob sich die Vpn in den beiden Zwischensubjektbedingungen (*Fremdeinschätzung* vs. *Selbsteinschätzung*) auf den Kovariaten unterscheiden. Wie durch die randomisierte Zuweisung angestrebt, zeigten sich keine signifikanten Unterschiede zwischen den Bedingungen in den Einstellungen zu Inklusion ( $t(131) = -0,39$ ,  $p > 0,05$ ), den Einstellungen zu Bildungsstandards ( $t(130) = 0,53$ ,  $p > 0,05$ ), den Erfahrungen mit Inklusion ( $t(132) = 0,41$ ,  $p > 0,05$ ), den Erfahrungen mit Bildungsstandards ( $t(128) = -0,06$ ,  $p > 0,05$ ) sowie der Selbstwirksamkeitserwartung bezogen auf inklusiven Unterricht ( $t(131) = -0,27$ ,  $p > 0,05$ ) und der Selbstwirksamkeitserwartung bezogen auf standardorientierten Unterricht ( $t(129) = 1,27$ ,  $p > 0,05$ ).

## 6.2 Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung

Hypothese 1a. besagte, dass sich aus Sicht von Lehrkräften die Unterrichtsgestaltung im inklusiven Unterricht von jener im standardorientierten Unterricht hinsichtlich zentraler Merkmale von Unterrichtsqualität unterscheidet. Erwartungskonform zeigte die einfaktorielle MANOVA einen signifikanten Haupteffekt des *Unterrichtskontexts* auf die Einschätzung der Unterrichtsgestaltung (Wilks  $\lambda = 0,53$ ,  $F(6,122) = 17,93$ ,  $p < 0,001$ , part.  $\eta^2 = 0,47$ ). Die Effektstärke fällt hoch aus. Univariat zeigte sich, dass auf allen Merkmalen von Unterrichtsqualität signifikante Unterschiede zwischen inklusivem und standardorientiertem Unterricht gesehen wurden, außer bzgl. Methodenvielfalt und Unterstützung durch die Lehrkraft (vgl. Tab. 2).

Tabelle 2

*Univariate Effekte des Innersubjektfaktors Unterrichtskontext auf die Merkmale von Unterrichtsqualität*

Abhängige Variable	<i>F</i>	<i>p</i>	part. $\eta^2$
Methodenvielfalt	0,00	0,956	0,00
Frontalunterricht	57,78	0,000	0,31
Differenzierung	14,32	0,000	0,10
Kognitive Aktivierung	45,97	0,000	0,27
Unterstützung durch die Lehrkraft	2,13	0,147	0,02
Klassenführung	7,83	0,006	0,06

Bonferroni-korrigierte post-hoc Tests erbrachten, dass mit inklusivem Unterricht weniger Frontalunterricht verbunden wurde als mit standardorientiertem (-0,64, 95%-

CI[-0,81, -0,47],  $p < 0,001$ ), eine geringer ausgeprägte kognitive Aktivierung (-0,44, 95%-CI[-0,57, -0,31],  $p < 0,001$ ), eine weniger effiziente Klassenführung (-0,18, 95%-CI[-0,30, -0,05],  $p < 0,05$ ) und eine stärker ausgeprägte Differenzierung (0,30, 95%-CI[0,14, 0,46],  $p < 0,001$ ).

Hypothese 1b. besagte, dass sich aus Sicht von Lehrkräften die Leistungsbeurteilung im inklusiven Unterricht von jener im standardorientierten Unterricht hinsichtlich der drei BNO unterscheidet. Erwartungskonform zeigte die einfaktorielle MANOVA einen signifikanten Haupteffekt des Unterrichtskontexts auf die Einschätzung der Leistungsbeurteilung (Wilks  $\lambda = 0,66$ ,  $F(3,129) = 21,86$ ,  $p < 0,001$ , part.  $\eta^2 = 0,34$ ). Die Effektstärke fällt ebenfalls hoch aus. Die univariate Betrachtung zeigte, dass auf allen drei BNO signifikante Unterschiede zwischen inklusivem und standardorientiertem Unterricht gesehen wurden (vgl. Tab. 3).

Tabelle 3

*Univariate Effekte des Innersubjektfaktors Unterrichtskontext auf die BNO*

Abhängige Variable	$F$	$p$	part. $\eta^2$
Individuelle BNO	14,60	0,000	0,10
Soziale BNO	56,38	0,000	0,30
Kriteriale BNO	47,95	0,000	0,27

Mit inklusivem Unterricht wurde eine schwächer ausgeprägte soziale BNO (-0,49, 95%-CI[-0,62,-0,36],  $p < 0,001$ ) und eine schwächer ausgeprägte kriteriale BNO (-0,43, 95%-CI[-0,55, -0,31],  $p < 0,001$ ) verbunden als mit standardorientiertem Unterricht, jedoch eine stärker ausgeprägte individuelle BNO (0,22, 95%-CI[0,11, 0,34],  $p < 0,001$ ).

Zusammengefasst bestätigte sich, dass sich aus Sicht von Lehrkräften die Unterrichtsgestaltung bzw. Leistungsbeurteilung im inklusiven Unterricht von jener im standardorientierten Unterricht unterscheidet. Es wurden unterschiedliche Ausprägungen von Frontalunterricht, Differenzierung, kognitiver Aktivierung, effizienter Klassenführung sowie individueller, sozialer und kriterialer BNO mit beiden Unterrichtskontexten verbunden, während das jeweilige Ausmaß an Methodenvielfalt und Unterstützung durch die Lehrkraft ähnlich eingeschätzt wurde.

Hypothese 2a. besagte, dass die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung auch unter Kontrolle der Einstellungen zu und der Erfahrungen mit Inklusion und Bildungsstandards und der Selbstwirksamkeitserwartungen bezogen auf inklusiven und standardorientierten Unterricht bestehen bleiben. Die einfaktorielle MANCOVA mit Einbezug dieser Kovariaten zeigte erneut einen signifikanten Haupteffekt des *Unterrichtskontexts* auf die Einschätzung der Unterrichtsgestaltung (Wilks  $\lambda=0,85$ ,  $F(6,108)=3,27$ ,  $p<0,05$ , part.  $\eta^2=0,15$ ). Die Effektstärke fällt hoch aus. Insgesamt blieben die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung auch unter Kontrolle der Kovariaten bestehen. Univariat zeigten sich signifikante Unterschiede jedoch nur noch bzgl. Frontalunterricht ( $F(1,108)=9,21$ ,  $p<0,05$ , part.  $\eta^2=0,08$ ), kognitiver Aktivierung ( $F(1,108)=6,97$ ,  $p<0,05$ , part.  $\eta^2=0,06$ ) und tendenziell Differenzierung ( $F(1,108)=3,74$ ,  $p=0,06$ , part.  $\eta^2=0,03$ ), wobei die Effektstärken klein bis mittelhoch ausfallen.

Hypothese 2b. beinhaltete, dass die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Leistungsbeurteilung ebenfalls unter Kontrolle der genannten Kovariaten bestehen bleiben. Die einfaktorielle MANCOVA zeigte keinen signifikanten Haupteffekt des *Unterrichtskontexts* auf die Einschätzung der verwendeten Bezugsnormen bei der Leistungsbeurteilung (Wilks  $\lambda=0,99$ ,  $F(3,116)=0,52$ ,  $p>0,05$ ). Erwartungswidrig gab es

unter Kontrolle der Kovariaten keine Unterschiede mehr zwischen den mit Inklusion und Standardorientierung verbundenen BNO.

Zusammenfassend konnte Hypothese 2a. insgesamt gestützt werden: Auch unter Berücksichtigung des Einflusses der Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie der Selbstwirksamkeitserwartung bzgl. inklusivem und standardorientiertem Unterricht zeigten sich Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung. Diese fanden sich jedoch univariat nur noch bzgl. Frontalunterricht, kognitiver Aktivierung und tendenziell Differenzierung. Erwartungswidrig wurde Hypothese 2b. nicht bestätigt. Bzgl. der drei BNO zeigten sich unter Kontrolle der Kovariaten keine signifikanten Unterschiede zwischen mit Inklusion und Bildungsstandards verbundener Leistungsbeurteilung.

### **6.3 Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung in Abhängigkeit von der Beurteilungsperspektive**

Hypothese 3a. beinhaltete, dass die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung je nach Beurteilungsperspektive variieren: Sie sollten bei der Fremdeinschätzung einer anderen Lehrkraft größer ausfallen als bei der Selbsteinschätzung.

Die 2 (*Unterrichtskontext*: Inklusion vs. Bildungsstandards) x 2 (*Beurteilungsperspektive*: Selbsteinschätzung vs. Fremdeinschätzung) faktorielle MANCOVA (Kovariaten: Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie Selbstwirksamkeitserwartungen bezogen auf inklusiven und standardorientierten Unterricht) zeigte einen signifikanten Haupteffekt des Unterrichtskontexts auf die Einschätzung der Unterrichtsgestaltung (Wilks  $\lambda = 0,85$ ,  $F(6,107) = 3,11$ ,  $p < 0,05$ , part.  $\eta^2 = 0,15$ ) (vgl. Hypothese 2a).

Dieser wurde durch einen Interaktionseffekt zwischen dem Innersubjektfaktor *Unterrichtskontext* mit dem Zwischensubjektfaktor *Beurteilungsperspektive* qualifiziert (Wilks  $\lambda=0,86$ ,  $F(6,107)=2,94$ ,  $p<0,05$ , part.  $\eta^2=0,14$ ). Die Effektstärke fällt groß aus.

Erwartungskonform variierten die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung je nach Beurteilungsperspektive.

Signifikante Interaktionseffekte zeigten sich univariat bzgl. Differenzierung ( $F(1,107)=8,06$ ,  $p<0,05$ , part.  $\eta^2=0,07$ ), Unterstützung durch die Lehrkraft ( $F(1,107)=13,74$ ,  $p<0,001$ , part.  $\eta^2=0,11$ ) und tendenziell effizienter Klassenführung ( $F(1,107)=3,71$ ,  $p=0,06$ , part.  $\eta^2=0,03$ ). Zur genaueren Analyse der Interaktionseffekte wurden getrennt nach Beurteilungsperspektive (*Selbsteinschätzung* vs. *Fremdeinschätzung*) einfaktorielles MANCOVA mit Bonferroni-korrigierten post-hoc Tests berechnet. Die zugehörigen Mittelwertmuster zeigt Abbildung 1.

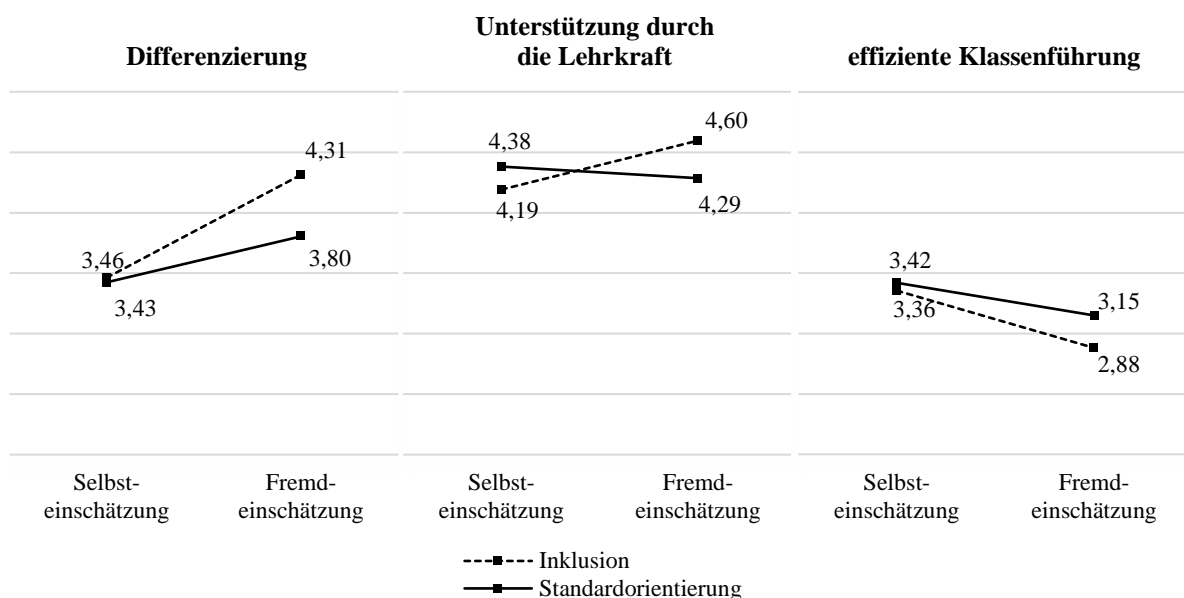


Abbildung 1. Differenzierung, Unterstützung durch die Lehrkraft und effiziente Klassenführung in Abhängigkeit von Unterrichtskontext und Beurteilungsperspektive

Es zeigte sich, dass in der Bedingung Fremdeinschätzung eine signifikant stärker ausgeprägte Differenzierung mit inklusivem Unterricht verbunden wurde als mit



standardorientiertem Unterricht (0,51, 95%-CI[0,33, 0,69],  $p < 0,001$ ). In der Bedingung Selbsteinschätzung fand sich jedoch kein Unterschied (0,03,  $p > 0,05$ ). Auch die Unterstützung durch die Lehrkraft wurde ausschließlich in der Bedingung Fremdeinschätzung als signifikant stärker im inklusiven als im standardorientierten Unterricht eingeschätzt (0,31, 95%-CI[0,21, 0,41],  $p < 0,001$ ), nicht aber in der Bedingung Selbsteinschätzung (-0,19,  $p > 0,05$ ). Die effiziente Klassenführung, für die sich tendenziell ein Interaktionseffekt zeigte, wurde ebenfalls lediglich in der Bedingung Fremdeinschätzung (-0,27, 95%-CI[-0,40, -0,13],  $p < 0,001$ ) als geringer ausgeprägt im inklusiven gegenüber standardorientiertem Unterricht gesehen, nicht aber in der Bedingung Selbsteinschätzung (-0,06,  $p > 0,05$ ).

Hypothese 3b. besagte, dass die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Leistungsbeurteilung bei der Fremdeinschätzung einer anderen Lehrkraft größer ausfallen, als bei der Selbsteinschätzung. Die 2 (*Unterrichtskontext*: Inklusion vs. Bildungsstandards) x 2 (*Beurteilungsperspektive*: Selbsteinschätzung vs. Fremdeinschätzung) faktorielle MANCOVA (Kovariaten: Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie Selbstwirksamkeitserwartungen bezogen auf inklusiven und standardorientierten Unterricht) zeigte zunächst keinen signifikanten Haupteffekt des *Unterrichtskontexts* auf die Einschätzung der Leistungsbeurteilung (Wilks  $\lambda = 0,99$ ,  $F(3,115) = 0,34$ ,  $p > 0,05$ ) (vgl. Hypothese 2b.). Jedoch zeigte sich ein signifikanter Interaktionseffekt zwischen dem Innersubjektfaktor *Unterrichtskontext* und dem Zwischensubjektfaktor *Beurteilungsperspektive* (Wilks  $\lambda = 0,93$ ,  $F(3,115) = 2,97$ ,  $p < 0,05$ , part.  $\eta^2 = 0,07$ ). Die Effektstärke fällt mittelhoch aus. Erwartungskonform variierten die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen BNO je nach Beurteilungsperspektive.

Univariat fanden sich signifikante Interaktionseffekte für die individuelle BNO ( $F(1,115)= 6,77, p < 0,05, \text{part. } \eta^2= 0,06$ ) und die kriteriale BNO ( $F(1,115)= 5,08, p < 0,05, \text{part. } \eta^2= 0,04$ ). Die Effektstärken fallen klein bis mittelhoch aus.

Zur genaueren Interpretation der Interaktionseffekte wurden erneut getrennt nach Beurteilungsperspektive (*Selbsteinschätzung* vs. *Fremdeinschätzung*) einfaktorielles MANCOVA mit Bonferroni-korrigierten post-hoc Tests berechnet. Die zugehörigen Mittelwertmuster zeigt Abbildung 2.

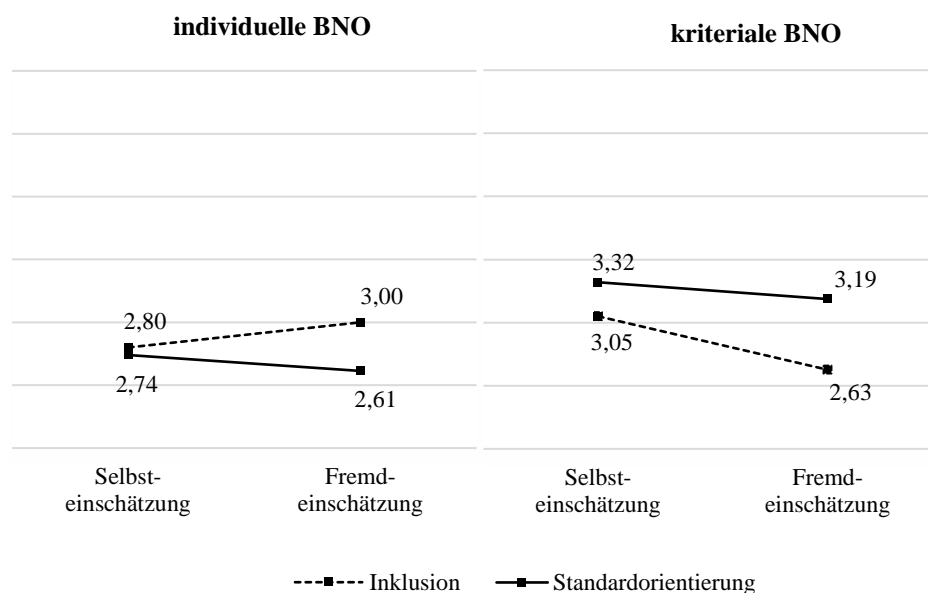


Abbildung 2. Individuelle und kriteriale BNO in Abhängigkeit von Unterrichtskontext und Beurteilungsperspektive

Nur in der Bedingung Fremdeinschätzung fand sich ein signifikanter Unterschied zwischen der mit inklusivem und standardorientiertem Unterricht verbundenen individuellen BNO ( $0,39, 95\text{-CI}[0,20, 0,57], p < 0,001$ ) zugunsten inklusiven Unterrichts, nicht aber in der Bedingung Selbsteinschätzung ( $0,06, p > 0,05$ ). Für die kriteriale BNO zeigte sich, dass diese sowohl in der Bedingung Fremdeinschätzung ( $-0,56, 95\text{-CI}[-0,73, -0,39], p < 0,001$ ) als auch in der Bedingung Selbsteinschätzung ( $-0,27, 95\text{-CI}[-0,45, -0,09], p < 0,05$ ) als

geringer ausgeprägt im inklusiven als im standardorientierten Unterricht eingeschätzt wurde. Die Mittelwertsdifferenz fällt jedoch in der Bedingung Fremdeinschätzung größer aus.

Hypothese 3a. und 3b. konnten insgesamt bestätigt werden. Die Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung variierten je nach Beurteilungsperspektive, jedoch nicht auf allen Merkmalen von Unterrichtsqualität und nicht auf allen BNO. Deutlicher als in den Hypothesen angenommen, gaben Lehrkräfte ausschließlich bei der Beurteilung des Unterrichts einer anderen Lehrkraft an, Unterschiede in der Differenzierung, Unterstützung durch die Lehrkraft, effizienten Klassenführung (tendenziell) und individuellen BNO mit beiden Unterrichtskontexten zu verbinden, nicht aber in der Selbsteinschätzung. Die kriteriale BNO wurde in beiden Beurteilungsperspektiven als weniger ausgeprägt im inklusivem gegenüber standardorientiertem Unterricht eingeschätzt, wobei der Unterschied bei der Fremdeinschätzung größer ausfiel.

## **7 Diskussion**

Unsere Studie untersuchte, inwiefern Lehrkräfte im Vorbereitungsdienst jeweils eine unterschiedliche Unterrichtsgestaltung und Leistungsbeurteilung mit Inklusion und Standardorientierung verbinden, wobei für die Implementation beider Reformansätze relevante Einflussgrößen und unterschiedliche Urteilspektiven berücksichtigt wurden.

Es zeigte sich, wie in Hypothese 1a./1b. angenommen, dass sich aus Sicht von Lehrkräften inklusiver Unterricht von standardorientiertem Unterricht hinsichtlich zentraler Merkmale von Unterrichtsqualität und Leistungsbeurteilung insgesamt unterscheidet. Im Einzelnen wurde mit inklusivem Unterricht weniger Frontalunterricht, eine geringer ausge-

prägte kognitive Aktivierung, eine weniger effiziente Klassenführung, eine geringer ausgeprägte soziale und kriteriale BNO, jedoch eine stärker ausgeprägte Differenzierung und individuelle BNO verbunden als mit standardorientiertem Unterricht. Somit konnte erstmals empirisch gezeigt werden, dass Lehrkräfte unterschiedliche Implikationen für Unterrichtsgestaltung mit Inklusion und Standardorientierung verbinden. Das Ausmaß an Methodenvielfalt und Unterstützung durch die Lehrkraft wurde jedoch in beiden Unterrichtskontexten ähnlich eingeschätzt.

Unter Kontrolle von für die Implementation beider Reformansätze relevanten Einflussgrößen (Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards und Selbstwirksamkeitserwartung bezogen auf inklusiven und standardorientierten Unterricht) wurden insgesamt ebenfalls Unterschiede in der Unterrichtsgestaltung im Kontext beider Reformansätze wahrgenommen, wodurch Hypothese 2a. bestätigt wurde. Die wahrgenommenen Unterschiede hinsichtlich Frontalunterricht, kognitiver Aktivierung und (tendenziell) Differenzierung blieben unter Kontrolle dieser Einflussgrößen bestehen, nicht jedoch hinsichtlich effizienter Klassenführung. Dieser scheint auf die Einstellungen, Erfahrungen und Selbstwirksamkeitserwartungen der Lehrkräfte rückführbar zu sein. Da frühere Studien zeigten, dass die Selbstwirksamkeit von Lehrkräften in Bezug auf Inklusion eher gering ausgeprägt ist (z.B. Gebhardt et al. 2015), sollte weitere Forschung die genauen Wirkmechanismen dieser Einflussgrößen mit Blick auf Unterrichtsgestaltung und besonders effiziente Klassenführung untersuchen, die zentral für inklusiven Unterricht erscheint (Emmer und Stough 2001).

Die Befunde deuten zudem darauf hin, dass sich Lehrkräfte der intendierten Wirkung von Bildungsstandards, kognitiv aktivierenden Unterricht anzuregen (Lenski et al. 2015) und jener von Inklusion, Differenzierung bzw. Individualisierung zu fördern und Frontalunterricht zu verringern (Kullmann et al. 2014) bewusst sind – und zwar unabhängig davon

ob sie den Reformen positiv oder negativ gegenüberstehen oder sich zutrauen, diese praktisch umzusetzen. Dass mit inklusivem Unterricht ein höheres Ausmaß an Differenzierung verbunden wurde, verweist auf den Diskurs darum, inwiefern ein im inklusiven Unterricht angemessener Grad an Differenzierung im Kontext von Standardorientierung erreicht werden kann (vgl. Seitz 2006). Wenngleich Richter et al. (2014) zeigten, dass Lehrkräfte, die VERA als Instrument zur Unterrichtsentwicklung begreifen, ihren Unterricht differenzierter gestalten, scheinen Lehrkräfte im Vorbereitungsdienst vergleichsweise wenig Differenzierung mit standardorientiertem Unterricht zu verbinden.

Erwartungswidrig wurde Hypothese 2b. nicht bestätigt: Auf den BNO wurden unter Kontrolle der Kovariaten keine Unterschiede zwischen beiden Reformansätzen gesehen. Die mit Inklusion und Bildungsstandards verbundenen Unterschiede in den BNO sind offenbar auf Einstellungen, Erfahrungen und Selbstwirksamkeitserwartungen rückführbar und die diskutierten unterschiedlichen Perspektiven auf Leistungsbeurteilung (Ainscow et al. 2006a) existieren nicht unabhängig von diesen. Dies steht im Einklang mit dem Befund, dass die BNO mit anderen Persönlichkeitsmerkmalen der Lehrkraft, wie z.B. Zielorientierungen, einhergeht (Schöne et al. 2004). Folglich ist denkbar, dass die BNO ein eher stabiles Merkmal ist (Rheinberg 1982) und weniger allein in Abhängigkeit vom jeweiligen Unterrichtskontext variiert.

Zudem zeigte sich, wie in Hypothese 3a./3b. angenommen, dass Unterschiede in der mit Inklusion und Bildungsstandards verbundenen Unterrichtsgestaltung und Leistungsbeurteilung je nach Beurteilungsperspektive variieren. Lehrkräfte gaben ausschließlich bei der Einschätzung des Unterrichts einer anderen Lehrkraft an, Unterschiede in der Differenzierung, Unterstützung durch die Lehrkraft, effizienten Klassenführung (tendenziell) und individuellen BNO mit beiden Unterrichtskontexten zu verbinden, nicht aber bzgl. des eigenen Unterrichts. Davon ausgehend, dass sich in der Beschreibung des Unterrichts einer anderen

Lehrkraft das jeweilige „Idealbild“ inklusiven bzw. standardorientierten Unterrichts widerspiegelt, nehmen Lehrkräfte folglich auf theoretischer Ebene unterschiedliche Implikationen beider Reformansätze wahr.

Für den Befund, dass in der Selbsteinschätzung keine unterschiedlichen Ausgestaltungen von Unterricht und Leistungsbeurteilung berichtet werden, sind verschiedene Erklärungen denkbar. Einerseits ist vorstellbar, dass Lehrkräfte nur auf theoretisch-konzeptioneller Ebene unterschiedliche Anforderungen mit Inklusion und Standardorientierung verbinden, nicht aber bezogen auf ihre eigene Praxis. Jedoch ist aufgrund früherer Studien denkbar, dass Lehrkräfte ihren eigenen Unterricht möglichst positiv beschreiben, was auf Verzerrungen in selbstwertdienliche bzw. sozial erwünschte Richtung (Wubbels et al. 1992; Clausen 2002) oder einen *self-presentation bias* (Kopcha und Sullivan 2007) rückführbar sein kann. Wird die gleichzeitige Umsetzung beider Reformen als wünschenswert bzw. notwendig angesehen, erscheint es positiver, keine Unterschiede in der eigenen Unterrichtsgestaltung und Leistungsbeurteilung zu berichten und „guten Unterricht“ unabhängig von beiden Reformen zu beschreiben. Allerdings ist zu betonen, dass die Befragten die eigene Unterrichtsgestaltung und Leistungsbeurteilung in beiden Unterrichtskontexten nicht grundsätzlich positiver beschrieben als die einer anderen Lehrkraft (vgl. Abb. 1 und Abb. 2). Dies könnte darauf zurückzuführen sein, dass es sich um Lehrkräfte im Vorbereitungsdienst handelte, deren Berufseinstiegsphase von Unsicherheit und Stress geprägt ist (vgl. Zimmermann und Klusmann 2016), was eine weniger positive Selbstbeschreibung nach sich zieht. Eine andere Erklärung ist, dass zwar unterschiedliche Anforderungen mit Inklusion und Standardorientierung verbunden werden, die Reformen aber die eigene Praxis kaum beeinflussen, was dem Befund entspricht, dass Unterricht nur schwer durch Reformen veränderbar ist (vgl. Gräsel und Parchmann 2004).

Zu berücksichtigen ist, dass sich unsere Befunde ausschließlich auf Berliner Lehrkräfte im Vorbereitungsdienst beziehen. Offen bleibt, inwiefern diese auf andere Stichproben, wie z.B. Lehrkräfte mit langjähriger Berufserfahrung, übertragbar sind. Weiterhin brachte unser experimentelles Design mit sich, dass die Vpn Inklusion und Standardorientierung als voneinander getrennte Unterrichtskontexte betrachteten. Damit ist eine Einschränkung der externen Validität verbunden, da eine solche Trennung in der Praxis nicht anzutreffen ist. Außerdem ist anzumerken, dass die Merkmale von Unterrichtsqualität jeweils über eine einzelne Skala erfasst wurden, wenngleich bspw. zu effizienter Klassenführung noch weitere Aspekte außer dem hier abgefragten Umgang mit Störungen gehören (z.B. Ophardt und Thiel 2017).

Zusammengefasst liefert unsere Studie erste Ergebnisse zu den unterschiedlichen Anforderungen an Unterrichtsgestaltung und Leistungsbeurteilung, die aus Sicht von Lehrkräften mit Inklusion und Standardorientierung verbunden sind. Diese sollten bei der Implementation beider gezielt aufgegriffen und die Lehrkräfte bei der bewussten Reflexion potentieller Gegensätzlichkeiten unterstützt werden, damit die Stärken beider Ansätze in der Praxis genutzt werden können. Eine offene Frage bleibt jedoch, inwiefern wahrgenommene Unterschiede aus Sicht von Lehrkräften letztlich zu widersprüchlichen Anforderungen oder „paradoxe[n] Handlungsanforderungen“ (Teumer 2012, o. S.) führen. Zukünftige Forschung sollte die tatsächliche Unterrichtsgestaltung im Kontext der Gleichzeitigkeit beider Reformen betrachten, um zu untersuchen, ob und inwiefern assoziierte Gegensätze Lehrkräfte vor Herausforderungen oder gar Dilemmata stellen und inwiefern z.B. positive Einstellungen und Selbstwirksamkeitserwartungen die subjektiv wahrgenommene Vereinbarkeit beider Reformen verbessern können.

## 8 Literatur

- Ainscow, M., Booth, T., & Dyson, A. (2006a). *Improving schools, developing inclusion*. London: Routledge.
- Ainscow, M., Booth, T., & Dyson, A. (2006b). Inclusion and the standards agenda. Negotiating policy pressures in England. *International Journal of Inclusive Education*, 10, 295–308.
- Beer, R. (2006). *Bildungsstandards. Einstellungen von Lehrerinnen und Lehrern*. Wien: Lit.
- Biewer, G. (2012). Die neue Welt der Bildungsstandards und ihre erziehungswissenschaftliche Rezeption aus der Perspektive einer Inklusiven Pädagogik. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 81, 9–21.
- Bober, M. J., Sullivan, H. J., Lowther, D. L., & Harrison, P. (1998). Instructional practices of teachers enrolled in educational technology and general education programs. *Educational Technology Research and Development*, 46(3), 81–97.
- Bosse, S., & Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empirische Sonderpädagogik*, 4, 279–299.
- Bräu, K. (2008). Die Betreuung selbstständigen Lernens. Vom Umgang mit Antinomien und Dilemmata. In G. Breidenstein & F. Schütze (Hrsg.), *Paradoxien in der Reform der Schule* (S. 179–199). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Davis, P., & Florian, L. (2004). *Teaching strategies and approaches for pupils with special educational needs: A scoping study*. DfES Research Report.
- Drieschner, E. (2009). *Bildungsstandards praktisch*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Emmer, E. T., & Stough, L. M. (2001). Classroom Management. A Critical Part of Educational Psychology, With Implications for Teacher Education. *Educational Psychologist*, 36, 103–112.
- Emmrich, R., Ernst, C.-M., Harych, P., & Wesselhöft, K. (2012). *Vergleichsarbeiten der Jahrgangsstufe 8 in Berlin als Beitrag zur Schul- und Unterrichtsentwicklung*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Eversheim, U. (2015). Bildungsstandards versus Inklusion? - Probleme und Potentiale der Standard- und Kompetenzorientierung für einen inklusiven (Sport)Unterricht. In S. Meier (Hrsg.), *Inklusion als Herausforderung. Aufgabe und Chance für den Schulsport* (S. 207–224). Berlin: Logos.
- Fischbach, A., Brunner, M., Krauss, S., & Baumert, J. (2015). Die Bezugsnormorientierung von Mathematiklehrkräften am Ende der Sekundarstufe I: Konvergenz verschiedener Messverfahren und Wirkung auf motivational-affektive Aspekte des Mathematiklernens und Leistung. *Journal for Educational Research Online*, 7(3), 3–27.



- Freudenthaler, H. H., & Specht, W. (2005). *Bildungsstandards aus Sicht der Anwender. Evaluation der Pilotphase I zur Umsetzung nationaler Bildungsstandards in der Sekundarstufe I*. Graz: Zentrum für Schulentwicklung.
- Gebhardt, M., Schwab, S., Nusser, L., & Hessels, M. G. P. (2015). Einstellungen und Selbstwirksamkeit von Lehrerinnen und Lehrern zur schulischen Inklusion in Deutschland - eine Analyse mit Daten des Nationalen Bildungspanels Deutschlands (NEPS). *Empirische Pädagogik*, 29, 211–229.
- Gräsel, C., & Parchmann, I. (2004). Implementationsforschung - oder: Der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 196–214.
- Grillitsch, M. (2010). *Bildungsstandards auf dem Weg in die Praxis. Ergebnisse einer Befragung von Lehrkräften und Schulleiter/innen der Sekundarstufe I zur Rezeption der Bildungsstandards und deren Implementation*. Graz: Leykam.
- Halbheer, U., & Reusser, K. (2008). Outputsteuerung, Accountability, Educational Governance. Einführung in Geschichte, Begrifflichkeiten und Funktionen von Bildungsstandards. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 26, 253–266.
- Helmke, A., & Hosenfeld, I. (2003). Vergleichsarbeiten (VERA): eine Standortbestimmung zur Sicherung schulischer Kompetenzen - Teil 1: Grundlagen, Ziele, Realisierung. *Schulverwaltung NRW, Ausgabe Hessen/Rheinland-Pfalz/Saarland(1)*, 10–13.
- Holder, K., & Kessels, U. (im Druck). Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung. *Unterrichtswissenschaft*.
- KMK (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters Kluwer.
- KMK (2011). *Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen. (Beschluss der Kultusministerkonferenz vom 20.10.2011)*. [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2011/2011\\_10\\_20-Inklusive-Bildung.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf). Zugegriffen: 21. Oktober 2015.
- Köller, O. (2005). Bezugsnormorientierung von Lehrkräften: Konzeptuelle Grundlagen, empirische Befunde und Ratschläge für praktisches Handeln. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendungen* (S. 189–202). Stuttgart: Kohlhammer.
- Kopcha, T. J., & Sullivan, H. (2007). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development*, 55, 627–646.
- Kullmann, H., Lütje-Klose, B., & Textor, A. (2014). Allgemeine Didaktik für inklusive Lerngruppen – fünf Leitprinzipien als Grundlage eines Bielefelder Ansatzes der inklusiven Didaktik. In B. Amrhein & M. Dziak-Mahler (Hrsg.), *Fachdidaktik inklusiv. Auf der Suche nach didaktischen Leitlinien für den Umgang mit Vielfalt in der Schule* (S. 89–107). Münster: Waxmann.
- Kunter, M., Baumert, J., Blum, W., & Neubrand, M. (Hrsg.) (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.

- Lenski, A., Richter, D., & Pant, H. (2015). Kompetenzorientierung im Unterricht aus der Perspektive von Lehrkräften und Schülerinnen und Schülern. *Zeitschrift für Pädagogik*, 61, 712–737.
- Ophardt, D., & Thiel, F. (2017). Klassenmanagement als Basisdimension der Unterrichtsqualität. In M. K. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 245–266). Wiesbaden: Springer Fachmedien.
- Oser, F. K., & Baeriswyl, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *AERA's Handbook of Research on Teaching* (S. 1031–1065). Washington: American Educational Research Association.
- PISA Konsortium Deutschland (Hrsg.) (2006). *PISA 2003. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Pöhlmann, C., Pant, H. A., Frenzel, J., Roppelt, A., & Köller, O. (2014). Auswirkungen einer Intervention auf die Auseinandersetzung und Arbeit mit Bildungsstandards bei Mathematik-Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 17, 113–133.
- Racherbäumer, K., & Kühn, S. M. (2013). Zentrale Prüfungen und individuelle Förderung. *Zeitschrift für Bildungsforschung*, 3, 27–45.
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Göttingen: Hogrefe.
- Rheinberg, F. (1982). Bezugsnormorientierung von Lehramtsanwärtern im Verlauf ihrer praktischen Ausbildung. In F. Rheinberg (Hrsg.), *Jahrbuch für Empirische Erziehungswissenschaft* (S. 235–248). Düsseldorf: Schwann.
- Richter, D., Böhme, K., Becker, M., Pant, H. A., & Stanat, P. (2014). Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten. Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen von Schülerinnen und Schülern. *Zeitschrift für Pädagogik*, 60, 225–244.
- Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2004). Zielorientierung und Bezugsnormorientierung. Zum Zusammenhang zweier Konzepte. *Zeitschrift für Pädagogische Psychologie*, 18, 93–99.
- Seitz, S. (2006). Inklusive Didaktik nach PISA. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 75, 192–199.
- Seitz, S., & Platte, A. (2006). Unterricht zwischen Vielfalt und Standardisierung. In A. Platte, S. Seitz, & K. Terfloth (Hrsg.), *Inklusive Bildungsprozesse* (S. 122–131). Bad Heilbrunn: Klinkhardt.
- Seitz, S., & Scheidt, K. (2012). Vom Reichtum inklusiven Unterrichts - Sechs Ressourcen zur Weiterentwicklung. *Zeitschrift für Inklusion*(1-2).
- Swanson, H. L. (1999). Instructional Components That Predict Treatment Outcomes for Students With Learning Disabilities. Support for a Combined Strategy and Direct Instruction Model. *Learning Disabilities Research and Practice*, 14(3), 129–140.
- Teumer, S. (2012). Paradoxe Handlungsanforderungen an LehrerInnen im Spannungsfeld der Neugestaltung des Schulanfangs. *Zeitschrift für Inklusion*(3).
- Wacker, A., Rohlf, C., & Kramer, J. (2013). Sind Bildungsstandards Innovationsimpulse für Unterricht und Leistungsbeurteilung? Ein Querschnittvergleich der

- Einschätzungen von Lehrerinnen und Lehrern zu zwei Messzeitpunkten. *Zeitschrift für Bildungsforschung*, 3, 119–136.
- Wagner, W., Helmke, A., & Rösner, E. (2009). *Materialien zur Bildungsforschung*. 25/1: *Deutsch Englisch Schülerleistungen International. Dokumentation der Erhebungsinstrumente für Schülerinnen und Schüler, Eltern und Lehrkräfte*. Frankfurt am Main: DIPF; GPPF.
- Werning, R. (2014). Stichwort: Schulische Inklusion. *Zeitschrift für Erziehungswissenschaft*, 17, 601–623.
- Wilhelm, M. (2009). *Integration in der Sekundarstufe I und II. Wie die Umsetzung im Fachunterricht gelingt*. Weinheim, Basel: Beltz.
- Wilhelm, M. (2012). Entwicklungsdidaktik als Antwort auf den Anspruch der Individualisierung in der inklusiven Schule. *Zeitschrift für Inklusion*(1).
- Wubbels, T., Brekelmans, M., & Hooymayers, H. P. (1992). Do teacher ideals distort the self-reports of their interpersonal behavior? *Teaching and Teacher Education*, 8, 47–58.
- Zimmermann, F., & Klusmann, U. (2016). Editorial. Burnout und Stress beim Übergang in den Lehrerberuf. *Psychologie in Erziehung und Unterricht*, 63(4), 241–243.

## **4.2 Studie 2: Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung**

Anmerkung:

Dieser Teil der Dissertation entspricht einem Manuskript, welches zur Publikation der Zeitschrift *Unterrichtswissenschaft* angenommen wurde.

Holder, K. & Kessels, U. (angenommen). Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung. *Unterrichtswissenschaft*.

**Die finale Publikation ist verfügbar bei Springer unter/**

**The final publication is available at [link.springer.com](https://link.springer.com) via**

**<https://doi.org/10.1007/s42010-018-0013-0>**

### **Zusammenfassung**

In deutschsprachigen Schulsystemen treffen aktuell die Reformansätze der Inklusion und der Bildungsstandards aufeinander. Im bildungswissenschaftlichen Diskurs wird darauf hingewiesen, dass diese u.a. bezüglich ihrer Perspektive auf Leistungen zueinander in Widerspruch stehen könnten: Im Kontext von Inklusion sollte die individuelle Bezugsnorm im Vordergrund stehen, bei den Bildungsstandards dagegen die kriteriale und soziale Bezugsnorm. Empirische Belege für diese postulierten Unterschiede fehlen jedoch. In der vorliegenden experimentellen Studie wurde variiert, ob Lehrkräfte im Vorbereitungsdienst ( $N=57$ ) sich vor der Beurteilung von Schülerleistungen entweder mit dem Konzept „Bildungsstandards“ oder mit „Inklusion“ beschäftigten. Die Ergebnisse zeigen, dass die Konfrontation mit Bildungsstandards in Interaktion mit der Einstellung zu diesen einen Einfluss auf die Bezugsnormorientierung hat. Je nachdem, ob die Standards abgelehnt oder befürwortet wurden, bewerteten Lehrkräfte im Vorbereitungsdienst stärker nach individueller oder nach sozialer Bezugsnorm. Die Konfrontation mit Inklusion beeinflusste die Bezugsnormorientierung hingegen nicht, weder bei positiver noch bei negativer Einstellung zu Inklusion. Die Ergebnisse werden in Bezug auf das potentielle Spannungsverhältnis von Inklusion und Standardorientierung diskutiert.

Schlüsselwörter: Bezugsnormorientierung, Bildungsstandards, Inklusion

### Abstract

In German-speaking countries, at present, two main reforms are simultaneously implemented in the educational system: inclusion and educational standards. Given these policies imply specific views on achievement and on evaluating student performance – inclusion implies an individual reference norm orientation, while criterion and social reference norms are emphasized by the standards-agenda – the issue of whether the two are compatible figures prominently in current scientific debate. Although it is widely discussed that inclusion and educational standards may have different impacts on teaching, only few studies have addressed this issue. The present study attempts to fill this gap. Using an experimental design and taking account of subjects' preexisting attitudes, it was investigated whether exposure to the concepts of "inclusion" vs. "educational standards" affected trainee teachers' ( $N=57$ ) reference norm orientations. Results show that teachers preferred either an individual reference norm or a social reference norm depending on their attitudes towards educational standards after exposure to this concept. Confronting teachers with inclusion had no effect on their reference norm orientations regardless of a priori attitudes towards inclusion. Results are discussed with regard to the current debate on possibly conflicting demands imposed on teachers by inclusion vs. standardization reforms.

Keywords: reference norm orientations, educational standards, inclusion

## 1 Einleitung

In deutschsprachigen Schulsystemen treffen derzeit mit der Implementation der Bildungsstandards einerseits sowie der Umsetzung inklusiver Beschulung andererseits zwei Reformsätze aufeinander, die die bildungspolitische und bildungswissenschaftliche Debatte maßgeblich bestimmen (Eversheim, 2015). Im bildungswissenschaftlichen Diskurs wird darauf hingewiesen, dass Bildungsstandards und Inklusion in verschiedenen Aspekten konfliktieren können (z.B. Seitz 2006). Das damit einhergehende potentielle Spannungsverhältnis zwischen der Vorgabe von Standards einerseits und individueller Förderung andererseits birgt verschiedene Herausforderungen für Lehrkräfte (z.B. Buschkühle, Duncker & Oswald, 2009). Offen bleibt jedoch, ob und inwiefern sich diese theoretisch postulierten Unterschiede auf das Handeln von Lehrenden auswirken.

Besonderer Fokus wird in diesem Diskurs auf die Gegensätze im Leistungsverständnis sowie in der Leistungserfassung innerhalb inklusiver und auf Standards ausgerichteter Unterrichtsprozesse gelegt (z.B. Ainscow, Booth & Dyson, 2006a). Während im Kontext von Inklusion die individuelle Bezugsnorm bei der Leistungsbewertung im Vordergrund stehe (z.B. Wilhelm, 2009), implizierten die Bildungsstandards eher die kriteriale und soziale Bezugsnorm (Heinrich, 2010; Helmke & Hosenfeld, 2003). Da sich die Bezugsnormorientierungen auf motivationale, affektive und kognitive Schülermerkmale auswirken können (zusammenfassend Köller, 2005), sind diese Annahmen zur differentiellen Passung der verschiedenen Bezugsnormen zu Bildungsstandards und Inklusion von großer Bedeutung.

Mit der Leistungsbewertung untersucht die vorliegende experimentelle Studie einen Teilaspekt des Lehrerhandels. Theoretisch liegt die Annahme zugrunde, dass Verhalten stets das Ergebnis einer Interaktion von Merkmalen der Situation und Merkmalen der Person ist (Lewin, 1936). Entsprechend wurde an einer Stichprobe von Lehrkräften im Vorbereitungsdienst ( $N=57$ ) getestet, ob die Auseinandersetzung mit den Konzepten „Bildungsstandards“

und „Inklusion“<sup>12</sup> in Abhängigkeit von den Einstellungen zu diesen zur Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbewertung führt. So kann mit unserer Studie erstmals überprüft werden, ob die Konzepte „Bildungsstandards“ und „Inklusion“ unterschiedliche Tendenzen im Lehrerhandeln anstoßen.

## 2 Theoretischer Hintergrund

### 2.1 Die (Un-)Vereinbarkeit von Inklusion und Standardorientierung

Das Verhältnis von Individualisierung und Standardisierung wird im aktuellen bildungswissenschaftlichen Diskurs teilweise als scheinbar unauflösbares Spannungsfeld beschrieben (Altrichter, 2011) und bekommt durch die gleichzeitige Implementation von Inklusion und der Bildungsstandards eine neue Aktualität (Teumer, 2012).

Innerhalb inklusiven Unterrichts sollen alle Lernenden mit ihren unterschiedlichen Voraussetzungen wertgeschätzt und eine „effektive, passgenaue und individuelle Förderung jedes einzelnen Kindes“ (Grosche, 2015, S. 26) erreicht werden. Dabei kommen Individualisierung und Differenzierung eine besondere Bedeutung zu (z.B. Prengel, 2014). Mit den Bildungsstandards werden hingegen vergleichbare Lernziele für alle Schülerinnen und Schüler beschrieben (Klieme et al., 2003). Sie legen fachspezifisch fest, welche Kompetenzen die Lernenden bis zu einer bestimmten Jahrgangsstufe erreichen sollen. Überprüft wird dies mit standardisierten Tests wie den Vergleichsarbeiten (VERA). Ein standardorientierter Unterricht zeichnet sich u.a. durch eine Kompetenz- und Ergebnisorientierung aus (KMK, 2010).

Das Aufeinandertreffen von Inklusion und Standards führt zu einer mitunter kontroversen Debatte (Altrichter, 2011). Vertreter und Vertreterinnen outputorientierter

---

<sup>12</sup> Inklusion wird als mehrdimensionales, uneinheitlich definiertes Konstrukt mit „konzeptuellen Unschärfen“ (Grosche, 2015, S. 20) diskutiert. Das Konzept von Inklusion, welches in dieser Studie zugrunde liegt, bezieht sich auf das gemeinsame Unterrichten von Kindern mit und ohne sonderpädagogischen Förderbedarf und die damit verbundene Steigerung der Heterogenität innerhalb der Schülerschaft.



Qualitätssicherung betonen, dass die Bildungsstandards und ihre Überprüfung nicht im Gegensatz zu Individualisierung stehen, sondern diese sogar wesentlich befördern (z.B. Oelkers & Reusser, 2008), weil sie durch ihre Orientierungs- und Diagnosefunktion Klarheit über Lernstände verschaffen, damit die Basis individueller Förderung bilden und Gestaltungsspielräume für individualisiertes Lernen öffnen. Aus Sicht der Sonder- und Inklusionspädagogik wird die Standardisierung von Lernergebnissen jedoch auch kritisch betrachtet (z.B. Sander, 2005): Bildungsstandards stünden Inklusion und damit verbundenen Individualisierungsbemühungen konträr gegenüber, da es vorrangig um das Erreichen einer Leistungsnorm statt um die bestmögliche Förderung aller Lernenden gehe (Eversheim, 2015).

Als ein zentraler Unterschied zwischen Inklusion und Bildungsstandards gelten die jeweiligen Perspektiven auf Leistung (Ainscow et al., 2006a). Während bei Inklusion das Augenmerk auf dem Lernprozess liege, fokussieren Bildungsstandards Lernergebnisse und widersprechen mit ihren punktuellen, normierten Leistungsmessungen entwicklungsorientierten Formen der Leistungsbewertung (Seitz, 2006). Für Lehrkräfte ergibt sich im Spannungsfeld von individueller Förderung und Standardorientierung, sowohl individuelle Entwicklungen wertzuschätzen als auch übergreifende Kriterien zur Leistungsbewertung heranzuziehen (z.B. Jantowski, 2014). Biewer (2012, S. 16) kommt auf Basis dieser potentiellen Konflikte sogar zu dem Schluss, dass durch Bildungsstandards „die Etablierung eines inklusiven Bildungssystems beeinträchtigt“ wird. Allerdings haben Pant et al. (2013) deutlich gemacht, dass die Standards nicht für Lernende gelten, die z.B. aufgrund eines Förderbedarfs zielfähig unterrichtet werden.

Forschungsergebnisse aus England zeigen, dass Lehrkräfte durchaus Spannungen zwischen Inklusion und der Standard-Agenda wahrnehmen und deren parallele Umsetzung kritisch beurteilen (z.B. Ainscow, Booth & Dyson, 2006b). Wie Hofmann (2011) betont,

existieren im deutschsprachigen Raum jedoch nur wenige Studien zur Wahrnehmung eines etwaigen Spannungsverhältnisses durch Lehrkräfte. Vereinzelt gibt es Hinweise, dass die beiden Konzepte einander konträr gegenüberstehen. So sehen 88% der zu Bildungsstandards befragten österreichischen Lehrkräfte keine Möglichkeit zur größeren Individualisierung im Unterricht mithilfe von Standards (Beer, 2006). Racherbäumer und Kühn (2013) zeigten, dass Lehrkräfte teilweise eine Unvereinbarkeit individueller Förderung und zentraler, standardisierter Prüfungen wahrnehmen. So wurde die Umsetzung differenzierenden Unterrichts im Kontext zentraler Prüfungen als eingeschränkt wahrgenommen (Racherbäumer & Kühn, 2013). Empirische Befunde zur Wahrnehmung unterschiedlicher Anforderungen an die Leistungsbewertung durch Inklusion und Bildungsstandards aus Lehrersicht liegen jedoch unserer Kenntnis nach nicht vor.

### **2.2 Bezugsnormorientierung bei der Leistungsbeurteilung**

Unter Bezugsnorm wird der Standard verstanden, mit dem eine zu bewertende Leistung verglichen wird (Rheinberg, 1998). Rheinberg (1980) unterscheidet drei verschiedene Bezugsnormen. Bei der *individuellen Bezugsnorm* wird eine gezeigte Leistung zu den Vorleistungen eines Lernenden in Beziehung gesetzt. Eine Leistungssteigerung wird positiv bewertet, ein Leistungsabfall negativ. Bei der *sozialen Bezugsnorm* wird die Leistung eines Lernenden mit Leistungen einer sozialen Vergleichsgruppe, z.B. der Schulklasse, verglichen. Eine überdurchschnittliche Leistung wird positiv bewertet, eine unterdurchschnittliche Leistung negativ. Bei der *kriterialen* (oder auch sachlichen) *Bezugsnorm* wird eine Leistung anhand eines Kriteriums beurteilt, welches zuvor extern oder intern definiert wurde (z.B. Anforderungen des Lehrplans). Die Tendenz zur Verwendung einer Bezugsnorm wird als Bezugsnormorientierung (BNO) bezeichnet (Rheinberg, 1998). Die BNO einer Lehrkraft wirken sich – wie zahlreiche Studien zeigen –

differentiell auf motivationale und kognitive Schülermerkmale aus, wobei hauptsächlich die individuelle BNO (IBNO) und soziale BNO (SBNO) untersucht wurden, die kriteriale BNO (KBNO) hingegen selten (zusammenfassend Köller, 2005). Weitgehend ungeklärt ist die Frage, aus welchem Grund Lehrkräfte unterschiedliche BNO bei der Leistungsbeurteilung zeigen (vgl. Köller, 2005).

*Merkmale der Lehrkräfte als Einflussfaktoren auf die BNO.* Als eine mögliche Ursache für unterschiedliche BNO wurden Merkmale der Lehrperson untersucht, wie z.B. Unterschiede in den Erziehungszielen (z.B. Mischo & Rheinberg, 1995) und Zielorientierungen (Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2004). Auch gehen die BNO mit unterschiedlichem Lehrverhalten einher, wie z.B. unterschiedlichen Individualisierungstendenzen im Unterricht (zusammenfassend Köller, 2005). Inwiefern sich Einstellungen auf die BNO auswirken, wurde bisher unserer Kenntnis nach nicht untersucht, obwohl Einstellungen und Überzeugungen zu wichtigen Aspekten der professionellen Kompetenz von Lehrenden gehören (Baumert & Kunter, 2006). Einstellungen spielen auch bei der Implementation von Inklusion (z.B. Kullmann, Lütje-Klose, Textor, Berard & Schitow, 2014) und Bildungsstandards (z.B. Grillitsch, 2010) eine entscheidende Rolle. Die Einstellungen zu Inklusion von Lehrkräften und Lehramtsstudierenden im deutschsprachigen Raum sind allgemein eher neutral bis positiv (z.B. Kessels, Erbring & Heiermann, 2014; Trumpp, Janz, Heyl & Seifried, 2014). Den Bildungsstandards stehen Lehrkräfte in Deutschland im Mittel grundsätzlich offen gegenüber, wünschen sich jedoch mehr Unterstützung im Umgang damit (Pant, Vock, Pöhlmann & Köller, 2008). Für Österreich zeigte Grillitsch (2010), dass ca. zwei Drittel der Befragten den Standards neutral gegenüberstehen und bei ca. einem Viertel negative Einstellungen überwiegen. Da Inklusion und Bildungsstandards mit unterschiedlichen Perspektiven auf Leistungsbewertung assoziiert werden, ist anzunehmen, dass auch Einstellungen zu diesen Reformen die BNO beeinflussen könnten.

*Merkmale der Beurteilungssituation als Einflussfaktoren auf die BNO.* Der Einfluss von Merkmalen der Situation und des Kontextes auf die BNO von Lehrkräften wurde in wenigen Studien untersucht. Wurden experimentell Bewertungsziele variiert, trat beim Ziel der Selektion häufiger die SBNO und beim Ziel der Förderung häufiger die IBNO auf (Wilbert & Gerdes, 2009). Auch ging die Einführung zentraler Abiturprüfungen in Hessen und Bremen mit einer Zunahme der KBNO einher (Holmeier, 2012). Dies deutet darauf hin, dass die BNO in Abhängigkeit des Kontextes sowie der „fundamentalen [...] politisch-gesellschaftlichen Ziele“ (Sacher, 2009, S. 88f.) variieren und unterschiedlich sinnvoll erscheinen (Holmeier, 2012). Entsprechend ist anzunehmen, dass die Auseinandersetzung mit Inklusion oder mit Bildungsstandards die BNO von Lehrenden beeinflussen könnte.

*Die Interaktion von Merkmalen der Person und der Situation.* Bisher wurden sowohl Merkmale der Lehrperson als auch situationelle Merkmale als Einflussfaktoren auf die BNO untersucht. Die sozialpsychologische Forschung hat vielfach bestätigt, dass gezeigtes Verhalten stets das Ergebnis einer Interaktion von Person und Situation ist (Richard, Bond & Stokes-Zoota, 2003). Es wird u.a. davon ausgegangen, dass die Interpretation einer Situation je nach Merkmalen einer Person unterschiedlich ausfallen kann. Übertragen auf die BNO ist denkbar, dass ein Kontext, der entweder Inklusion oder Bildungsstandards hervorhebt, je nach Einstellungen zu diesen Reformen damit kongruentes oder aber entgegengesetztes Beurteilungsverhalten anstoßen kann.

### **3 Fragestellung und Hypothesen**

Zwischen Inklusion und Standardorientierung deuten sich in Bezug auf die Leistungsbewertung mögliche Widersprüche an. Inklusion impliziert die vorrangige Orientierung an der individuellen Bezugsnorm, da prozessorientierte Verfahren der Leistungsbewertung fokus-

siert werden (z.B. Seitz, 2006) und „jeder Entwicklungsfortschritt [...] als Leistung bewertet“ (Wilhelm, 2009, S. 107) werden soll. Bildungsstandards hingegen betonen die kriteriale und soziale Bezugsnorm (Helmke & Hosenfeld, 2003), da hier Leistungen an einem kriterialen Maßstab gemessen und in den Ergebnisrückmeldungen zu VERA üblicherweise zu einer sozialen Bezugsnorm (klassenintern, klassen- oder schulübergreifend) in Beziehung gesetzt werden (Helmke & Hosenfeld, 2003). Zu den Auswirkungen dieser potentiellen Gegensätze auf die Leistungsbeurteilung liegen jedoch unserer Kenntnis nach noch keine Studien vor. Zudem fokussierten die wenigen vorliegenden Studien zu den Einflussfaktoren auf die BNO bisher entweder Merkmale der Person oder Merkmale der Situation, ohne einer Wechselwirkung beider Rechnung zu tragen.

In der vorliegenden experimentellen Studie soll daher überprüft werden, ob die Auseinandersetzung mit dem Thema Bildungsstandards oder mit dem Thema Inklusion je nach Einstellung zu diesen Konzepten zu unterschiedlichen Vorgehensweisen bei der Leistungsbewertung führt. Aus der sozialpsychologischen Forschung ist bekannt, dass Einstellungen durch die Konfrontation mit den Einstellungsobjekten zugänglich gemacht und somit handlungswirksam werden (z.B. Fazio, Chen, McDonel & Sherman, 1982). Deshalb nehmen wir an, dass bei Personen mit einer positiven Einstellung zu Bildungsstandards eine Konfrontation mit dem Thema Bildungsstandards zu einer geringeren Ausprägung der IBNO und zu einer stärkeren Ausprägung der KBNO und der SBNO führt. Werden die Standards eher abgelehnt, sollte durch die Auseinandersetzung damit diese negative Einstellung handlungsleitend werden und entsprechend zu einer deutlich stärkeren IBNO und einer geringeren KBNO und SBNO führen. Ein spiegelbildliches Muster wird bei der Konfrontation mit dem Thema Inklusion in Abhängigkeit von der zugehörigen Einstellung erwartet. In einem experimentellen Design, in dem variiert wird, ob sich die

Teilnehmenden entweder mit dem Thema Bildungsstandards oder mit dem Thema Inklusion beschäftigen, werden entsprechend folgende Hypothesen geprüft:

1. Die Konfrontation mit Bildungsstandards führt bei positiven Einstellungen zu diesem Konzept zu einer stärkeren Ausprägung der KBNO und der SBNO und einer schwächeren Ausprägung der IBNO.
2. Die Konfrontation mit Bildungsstandards führt bei negativen Einstellungen zu diesem Konzept zu einer stärkeren Ausprägung der IBNO und einer schwächeren Ausprägung der KBNO und der SBNO.
3. Die Konfrontation mit Inklusion führt bei positiven Einstellungen zu diesem Konzept zu einer stärkeren Ausprägung der IBNO und einer schwächeren Ausprägung der KBNO und der SBNO.
4. Die Konfrontation mit Inklusion führt bei negativen Einstellungen zu diesem Konzept zu einer stärkeren Ausprägung der KBNO und der SBNO und einer schwächeren Ausprägung der IBNO.

## 4 Methode

### 4.1 Design

Die Hypothesen wurden in einem Experiment überprüft. Die Versuchspersonen (VP) wurden randomisiert einer der beiden Untersuchungsbedingungen zugewiesen (*Konfrontation mit Inklusion* vs. *Konfrontation mit Bildungsstandards*). Die Einstellung zu diesen Konzepten wurde als quasi-experimenteller Faktor in die Auswertung einbezogen.

### 4.2 Stichprobe

Die Stichprobe umfasste  $N=94$  Berliner Lehrkräften im Vorbereitungsdienst. Um eine Verzerrung der Ergebnisse zu vermeiden, wurden bei der Auswertung alle VP ausgeschlossen, die angaben, die zur Messung der BNO verwendete „Kleine

Beurteilungsaufgabe“ (Rheinberg, 1980) bereits zu kennen (siehe Kapitel 4.3). Damit reduzierte sich die Stichprobe für die Datenanalyse auf  $N=57$  (Alter  $M=33.09$ ,  $SD=6.37$ ; 67% weiblich), wovon  $n=28$  der Bedingung *Konfrontation mit Inklusion* und  $n=29$  der Bedingung *Konfrontation mit Bildungsstandards* zugewiesen worden waren.

### 4.3 Untersuchungsinstrument und Durchführung

Die VP nahmen im Rahmen eines Seminars freiwillig an einer schriftlichen Befragung zum Thema „Neue Herausforderungen im Lehrerberuf“ teil, in die die im Folgenden beschriebenen Stimuli und Messinstrumente eingebettet waren.

Den VP wurde zuerst als experimenteller Stimulus ein Text präsentiert, in dem entweder das von uns zugrunde gelegte Konzept von Inklusion oder von Bildungsstandards überblicksartig dargestellt wurde. Der jeweilige Text wurde auf der Basis einschlägiger Publikationen entwickelt. Der Text zum Thema Bildungsstandards enthielt Informationen zu Zielen der Bildungsstandards, zu VERA sowie zu kompetenz- und ergebnisorientiertem Unterricht. Der Text zum Thema Inklusion informierte über Ziele inklusiver Pädagogik, individuelle Förderung sowie individualisierten Unterricht. Um sicherzustellen, dass sich die VP mit den Inhalten des jeweiligen Ansatzes beschäftigten, hatten sie die Aufgabe, die zentralen Inhalte des Textes stichpunktartig wiederzugeben.

Nach dieser Konfrontation mit Bildungsstandards oder Inklusion wurde als abhängige Variable die BNO über eine leicht modifizierte Version der „Kleinen Beurteilungsaufgabe“ (KBA) von Rheinberg (1980) gemessen, die ein verhaltensnahes Maß zur Erfassung der BNO darstellt (Dickhäuser & Rheinberg, 2003). Bei der KBA sind neun fiktive Schülerinnen und Schüler anhand ihrer Leistungen, die sie in drei zurückliegenden Tests erbracht haben, zu beurteilen. Drei der Kinder erzielten im letzten Test niedrige Punktzahlen (25 Punkte), drei mittlere (50 Punkte) und drei hohe Punktzahlen (75 Punkte). Die drei Kinder

mit dem gleichen Leistungsstand differieren in ihrer Leistungsentwicklung über die Tests hinweg: ein Kind zeigte eine Verbesserung, eines eine Verschlechterung und eines stagnierende Leistungen.

Die VP hatten die Leistungen auf einer Intervallskala zu bewerten (-5 Punkte bis +5 Punkte). Nach Rheinberg (1980) äußert sich die IBNO in einer positiven Beurteilung aller Kinder, die sich über die drei Tests hinweg verbesserten, ungeachtet dessen, wie diese verglichen mit den anderen abschnitten. Eine Bewertung unter SBNO berücksichtigt dagegen die individuelle Entwicklung nicht, sondern nur, wie gut die im letzten Test erreichte Leistung verglichen mit den Leistungen der anderen Kinder ausfällt. Mit der KBA ist eine Messung der KBNO nicht intendiert, weshalb diese für unsere Studie leicht modifiziert wurde. Über die Formulierung Rheinbergs (1980) hinausgehend, wurde bei der Aufgabenstellung der kriteriale Maßstab stärker hervorgehoben, indem erläutert wurde, dass minimal 0 Punkte und maximal 100 Punkte erreicht werden konnten. Eine Punktzahl von 75 spiegelt somit bspw. wider, dass 75% des vorgegebenen Kriteriums erreicht wurden. Die KBNO würde sich also in einem spezifischen Bewertungsmuster zeigen, in dem keinem der Kinder die schlechteste oder beste Bewertung (-5 Punkte bzw. +5 Punkte) zuteilwird, da keines die mögliche Minimal- oder Maximalpunktzahl erreichte. In Tabelle 1 ist die Beurteilungsskala, welche den VP zur Verfügung stand, zur vorgegeben kriterialen Punkteskala von 0 bis 100 in Beziehung gesetzt.



Tabelle 1

*Gegenüberstellung der Skala zur Beurteilung der fiktiven Testergebnisse und des vorgegebenen kriterialen Vergleichsmaßstabs*

Beurteilungsskala für Testergebnisse	kriterialer Vergleichsmaßstab
+++++	100
++++	90
+++	80
++	70
+	60
neutral	50
-	40
--	30
---	20
----	10
-----	0

Unter Verwendung der kriterialen Bezugsnorm sollten die Kinder, welche 75 Punkte im letzten Test erreichten, mit zwei oder drei Pluspunkten bewertet werden, jene mit 50 Punkten neutral und jene mit 25 Punkten mit zwei oder drei Minuspunkten. Nur VP, die genau dieses Bewertungsmuster zeigen, wurde die KBNO attestiert, da dabei weder die individuelle Entwicklung noch das Abschneiden relativ zu den anderen Kindern berücksichtigt werden.

Die Einstellungen zu Inklusion wurden über die Subskala „*Core Perspectives*“ der deutschen Version des „*My Thinking About Inclusion*“ (MTAI) Fragebogens (Paulus, 2013;

Stoiber, Gettinger & Goetz, 1998) erfasst. Daraus wurden fünf Items mit fünffach gestuftem Antwortformat ausgewählt, welche die Grundeinstellung zu Inklusion abbilden und z.B. Aussagen wie „Kinder mit besonderem Unterstützungsbedarf haben das Recht, im selben Klassenzimmer wie alle anderen unterrichtet zu werden“ enthalten. Es wurde ein Mittelwert über diese fünf Items berechnet ( $\alpha=.71$ ).

Die Einstellungen zu Bildungsstandards wurden über sechs ausgewählte Items der Skala „Einstellungen zu Bildungsstandards im Allgemeinen“ (Freudenthaler & Specht, 2005) erfasst. Es wurden jeweils drei Items mit fünffach gestuftem Antwortformat ausgewählt, welche positive (z.B. „Bildungsstandards geben Klarheit, worauf es in der Schule letztendlich ankommt“) und negative Aspekte (z.B. „Bildungsstandards bringen die Gefahr einer inhaltlichen Verarmung des Unterrichts mit sich“) der Implementation von Bildungsstandards abbilden. Letztere wurden vor der Bildung des Mittelwertes umgepolt, so dass ein höherer Wert einer positiveren Einstellung entspricht ( $\alpha=.78$ ). Die VP wurden nach Beendigung der Studie über deren Ziele und Ergebnisse aufgeklärt.

#### **4.4 Datenanalyse**

Zur Auswertung der KBA hat Rheinberg (1980) ein Verfahren vorgeschlagen, mit dem Urteilsdifferenzen zwischen Schülerpaaren mit unterschiedlichen Leistungsständen bzw. unterschiedlicher Leistungsentwicklung berechnet werden, um die BNO abzubilden. Dies hat jedoch den Nachteil, dass die Maßzahlen für die SBNO und die IBNO unterschiedliche Wertebereiche aufweisen und erst nach Transformation in eine gemeinsame Skala miteinander vergleichbar sind. Um für diese beiden BNO ein einheitliches, metrisch interpretierbares Maß zu erhalten, wurden für jede VP zwei Korrelationskoeffizienten berechnet, die das Ausmaß der Orientierung an der sozialen ( $r_{SBNO}$ ) und der individuellen Bezugsnorm ( $r_{IBNO}$ ) wiedergeben.  $r_{SBNO}$  wurde berechnet, indem die Leistungsbeurteilung jeder VP mit

der „idealtypischen Urteilskonfiguration“ (Rheinberg, 1980, S. 27) unter ausschließlicher Anwendung der sozialen Bezugsnorm korreliert wurde. Je näher der Koeffizient  $r_{SBNO}$  an Eins liegt, desto stärker ist die SBNO.  $r_{IBNO}$  wurde analog mit dem idealtypischen Bewertungsmuster unter IBNO berechnet. Je näher dieser Koeffizient an Eins liegt, desto stärker ist die IBNO. Die KBNO wurde über die Kodierung von Personen identifiziert, die ausschließlich das dafür charakteristische, oben erläuterte Bewertungsmuster zeigten (vgl. Tabelle 1). Da ein Abweichen von diesem Muster die Anwendung der KBNO ausschließt, wurde hier rein dichotom zwischen dem Vorhandensein (KBNO=1) und dem Nichtvorhandensein (KBNO=0) der KBNO unterschieden. Zur Überprüfung der Hypothesen wurden multiple Regressionsanalysen berechnet.

## 5 Ergebnisse

### 5.1 Deskriptive Ergebnisse

Lediglich eine Person wendete die KBNO bei der Beurteilung an. Aufgrund dieses Befundes können die auf die KBNO bezogenen Teilhypothesen nicht überprüft werden. Die IBNO und die SBNO ließen sich in der Stichprobe jedoch gut abbilden. Es zeigte sich, dass die IBNO bei den Befragten insgesamt weniger stark ausgeprägt war als die SBNO ( $Mr_{IBNO}=.57$ ;  $Mr_{SBNO}=.85$ ). Im Mittel zeigte sich kein signifikanter Unterschied in der Ausprägung der IBNO in der experimentellen Bedingung „Konfrontation mit Bildungsstandards“ ( $Mr_{IBNO}=.57$ ) und in der Bedingung „Konfrontation mit Inklusion“ ( $Mr_{IBNO}=.57$ ). In der Ausprägung der SBNO in der Bedingung „Konfrontation mit Bildungsstandards“ ( $Mr_{SBNO}=.86$ ) und in der Bedingung „Konfrontation mit Inklusion“ ( $Mr_{SBNO}=.84$ ) zeigte sich ebenfalls kein signifikanter Unterschied.

Die Einstellungen zu Bildungsstandards lagen im mittleren Bereich ( $M=2.87$ ;  $SD=.67$ ), während die Einstellungen zu Inklusion moderat positiv ausfielen ( $M=3.54$ ;

$SD=.70$ ). Beide Einstellungen waren unabhängig voneinander,  $r(55) = -.04$ ,  $p = .79$ . Es zeigte sich kein signifikanter Unterschied zwischen den Einstellungen zu Bildungsstandards in der experimentellen Bedingung „Konfrontation mit Bildungsstandards“ ( $M=2.89$ ;  $SD=.59$ ) und in der Bedingung „Konfrontation mit Inklusion“ ( $M=2.85$ ;  $SD=.77$ ). Ebenso zeigte sich kein signifikanter Unterschied zwischen den Einstellungen zu Inklusion in der Bedingung „Konfrontation mit Bildungsstandards“ ( $M=3.55$ ;  $SD=.61$ ) und in der Bedingung „Konfrontation mit Inklusion“ ( $M=3.53$ ;  $SD=.81$ ).

### **5.2 Effekte der Konfrontation mit Bildungsstandards auf die BNO in Abhängigkeit von der Einstellung zu Bildungsstandards**

Um zu prüfen, ob sich die Konfrontation mit Bildungsstandards je nach Einstellung zu diesen auf die BNO auswirkt (Hypothesen 1 und 2), wurden zwei multiple lineare Regressionsmodelle zur Vorhersage der IBNO einerseits und der SBNO andererseits berechnet. Zuvor wurde mithilfe der Varianzinflationsfaktoren (VIF) sichergestellt, dass keine Multikollinearität vorliegt (höchster VIF=1.61).

Als Prädiktoren gingen die Konfrontation mit dem Konzept der Bildungsstandards vs. dem Konzept Inklusion (als Dummyvariable) und die Einstellung zu Bildungsstandards sowie deren Interaktion simultan in die Berechnungen ein. Zur besseren Interpretation der Interaktionseffekte erfolgte eine Zentrierung der Einstellungsvariablen am Mittelwert. Die Ergebnisse für die Vorhersage der IBNO sind in Tabelle 2 dargestellt.

Tabelle 2

*Vorhersage der IBNO durch die Einstellung zu Bildungsstandards und die experimentelle Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)*

	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	<i>p</i>
Einstellung zu Bildungsstandards	-.06	.09	-.11	-.71	.48
Konfrontation mit Bildungsstandards vs. Inklusion (0=Inklusion)	.01	.09	.02	.14	.89
Einstellung Bildungsstandards x Konfrontation	-.30	.14	-.33	-2.10	.04

$N=57$ .  $R^2=.17$  ( $p=.02$ ).

Weder die Einstellung zu Bildungsstandards als solche noch die Konfrontation mit Bildungsstandards waren signifikante Prädiktoren für die IBNO. Erwartungsgemäß zeigte sich jedoch eine signifikante Wechselwirkung zwischen der Einstellung zu Bildungsstandards und der experimentellen Bedingung ( $\beta = -.33$ ,  $t(54) = -2.10$ ,  $p < .05$ ). Das zugrundeliegende Mittelwertmuster ist in Abbildung 1 dargestellt.

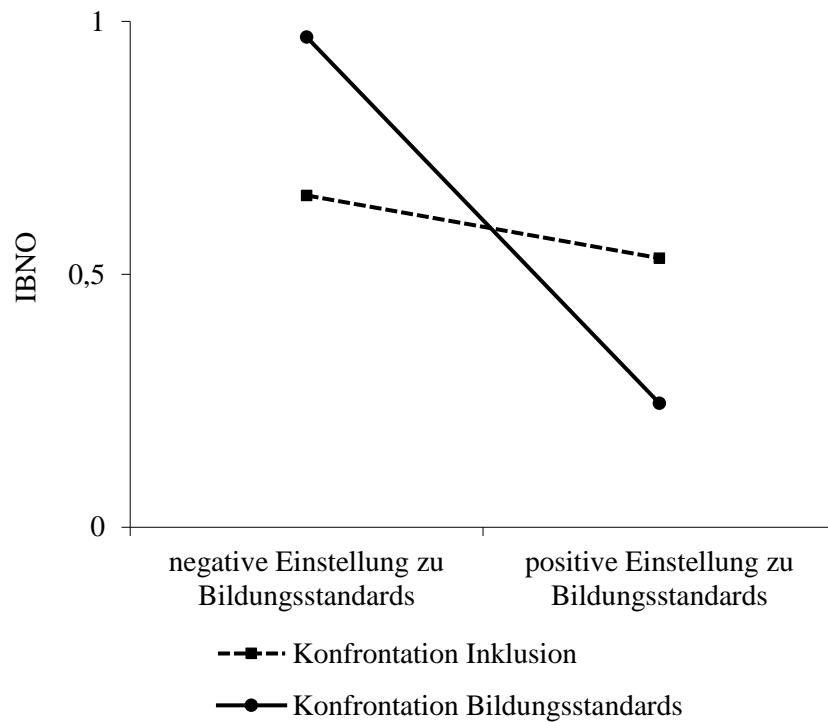


Abbildung 1. IBNO in Abhängigkeit von der Einstellung zu Bildungsstandards und der experimentellen Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)

Je positiver die Einstellung zu Bildungsstandards war, desto schwächer fiel die IBNO aus, wenn eine Konfrontation mit Bildungsstandards erfolgte. Je negativer eine Person gegenüber Bildungsstandards eingestellt war, desto stärker war die IBNO, wenn eine Konfrontation mit Bildungsstandards stattfand. Bei einer Konfrontation mit Inklusion zeigten sich ähnliche Ausprägungen der IBNO unabhängig von der Einstellung zu Bildungsstandards. Die Ergebnisse der Regressionsanalyse zur Vorhersage der SBNO zeigt Tabelle 3.

Tabelle 3

*Vorhersage der SBNO durch die Einstellung zu Bildungsstandards und die experimentelle Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)*

	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	<i>p</i>
Einstellung zu Bildungsstandards	.01	.08	.02	.12	.91
Konfrontation mit Bildungsstandards vs. Inklusion (0=Inklusion)	-.01	.09	-.01	-.11	.92
Einstellung Bildungsstandards x Konfrontation	.29	.14	.35	2.13	.04

$N=57$ .  $R^2=.13$  ( $p=.06$ ).

Das Regressionsmodell verfehlt mit einer Irrtumswahrscheinlichkeit von  $p=.06$  knapp das Signifikanzniveau von 5%. Ähnlich wie bei der IBNO beeinflussten weder die Einstellung zu Bildungsstandards als solche noch die Konfrontation damit für sich genommen das Ausmaß der SBNO. Erwartungskonform zeigte sich aber ebenfalls eine signifikante Interaktion zwischen der Einstellung zu Bildungsstandards und der Untersuchungsbedingung ( $\beta=.35$ ,  $t(54)=-2.13$ ,  $p<.05$ ), wie in Abbildung 2 dargestellt.

Je positiver die Einstellung zu Bildungsstandards war, desto stärker fiel die Ausprägung der SBNO aus, wenn eine Konfrontation mit Bildungsstandards stattfand. Je negativer die Einstellung zu Bildungsstandards ausfiel, desto schwächer war die SBNO ausgeprägt, wenn eine Konfrontation mit diesem Konzept erfolgte. Bei einer Auseinandersetzung mit Inklusion zeigten sich erneut ähnliche Ausprägungen der SBNO unabhängig von der Einstellung zu den Bildungsstandards.

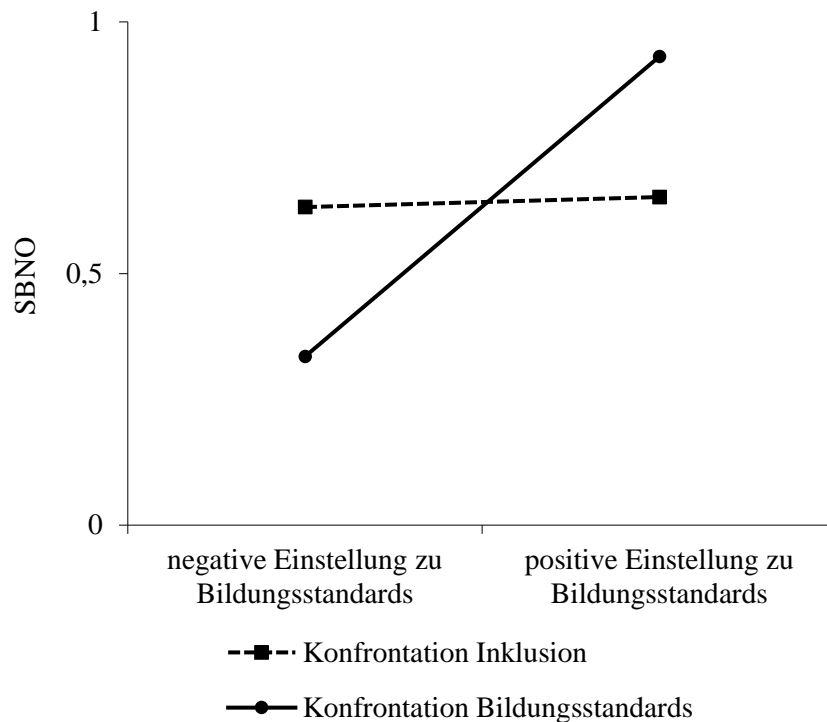


Abbildung 2. SBNO in Abhängigkeit von der Einstellung zu Bildungsstandards und der experimentellen Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)

Zusammenfassend kann Hypothese 1 bestätigt und Hypothese 2 tendenziell bestätigt werden: Bei Personen mit einer positiven Einstellung zu Bildungsstandards zeigte sich nach der Konfrontation damit eine geringere IBNO, wohingegen Personen mit einer negativen Einstellung zu Bildungsstandards eine stärkere IBNO zeigten, nachdem sie damit konfrontiert wurden. Hinsichtlich der SBNO zeigte sich tendenziell ein korrespondierendes Muster: Bei Personen mit einer positiven Einstellung zu Bildungsstandards führte eine Konfrontation mit diesem Thema zu einer verstärkten SBNO, wohingegen Personen mit einer negativen Einstellung zu Bildungsstandards eine schwächere SBNO zeigten.

### 5.3 Effekte der Konfrontation mit Inklusion auf die BNO in Abhängigkeit von der Einstellung zu Inklusion

Um zu prüfen, ob sich die Konfrontation mit Inklusion ja nach Einstellung zu diesem Konzept auf die BNO auswirkt (Hypothesen 3 und 4), wurde analog vorgegangen. Zuvor



wurde sichergestellt, dass keine Multikollinearität vorliegt (höchster VIF=1.59). Die Ergebnisse für die Vorhersage der IBNO durch die Einstellung zu Inklusion und die Konfrontation mit Inklusion vs. Bildungsstandards zeigt Tabelle 4.

Tabelle 4

*Vorhersage der IBNO durch die Einstellung zu Inklusion und die experimentelle Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)*

	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	<i>p</i>
Einstellung zu Inklusion	-.02	.09	-.04	-.20	.84
Konfrontation mit Bildungsstandards vs. Inklusion (0=Inklusion)	-.02	.10	-.03	-.19	.85
Einstellung Inklusion x Konfrontation	-.05	.15	-.06	-.34	.73

$N=55$ .  $R^2=.01$  ( $p=.94$ ).

Weder die Einstellung zu Inklusion als solche noch die Konfrontation damit waren signifikante Prädiktoren für die IBNO. Erwartungswidrig zeigte sich keine signifikante Interaktion zwischen der Einstellung zu Inklusion und der experimentellen Bedingung. Bei der Vorhersage des Ausmaßes der SBNO fanden sich ebenfalls keine signifikanten Effekte, wie in Tabelle 5 dargestellt.

Tabelle 5

*Vorhersage der SBNO durch die Einstellung zu Inklusion und die experimentelle Bedingung (Konfrontation mit Bildungsstandards vs. Inklusion)*

	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	<i>p</i>
Einstellung zu Inklusion	.02	.09	.04	.25	.80
Konfrontation mit Bildungsstandards vs. Inklusion (0=Inklusion)	.01	.10	.02	.15	.89
Einstellung Inklusion x Konfrontation	.06	.14	.07	.42	.67

$N=55$ .  $R^2=.01$  ( $p=.89$ ).

Zusammenfassend konnten Hypothese 3 und Hypothese 4 nicht bestätigt werden. Weder bei Personen, die positiv gegenüber Inklusion eingestellt sind, noch bei Personen mit einer negativen Einstellung zu Inklusion wirkt sich die Konfrontation damit auf die BNO aus.

## 6 Diskussion

Die vorliegende Studie untersuchte, ob mit Bildungsstandards und Inklusion in der Wahrnehmung von Lehrkräften im Vorbereitungsdienst unterschiedliche Implikationen für die Leistungsbeurteilung verbunden sind. Es wurde überprüft, inwiefern die Auseinandersetzung mit „Bildungsstandards“ oder „Inklusion“ in Abhängigkeit von der Einstellung zu diesen mit unterschiedlichen BNO einhergeht. Wie erwartet, zeigten Personen mit einer positiven Einstellung zu Bildungsstandards nach der Konfrontation mit diesem Konzept eine geringere IBNO. Personen mit einer negativen Einstellung zu Bildungsstandards zeigten hingegen eine stärkere IBNO, nachdem sie mit diesem Konzept konfrontiert wurden. Erwartungskonform zeigte sich tendenziell ein korrespondierendes Muster bei der SBNO. Bei Personen mit einer positiven Einstellung zu Bildungsstandards führte eine Konfrontation mit diesem Thema zu einer verstärkten SBNO, wohingegen Personen mit einer negativen

Einstellung zu Bildungsstandards eine schwächere SBNO zeigten. Entgegen der Erwartungen wirkte sich die Konfrontation mit Inklusion weder bei den Befragten mit einer positiven Einstellung noch bei jenen mit einer negativen Einstellung zu Inklusion auf die BNO aus. Eine mögliche Erklärung dafür ist, dass Inklusion mehrheitlich als „politisch korrekte“ Reform angesehen wird (Avramidis & Norwich, 2002) und deshalb die Beschäftigung damit kaum Positionierungen dafür oder dagegen auslöst, die sich in besser oder schlechter zur Inklusion passenden BNO äußern könnten. Bildungsstandards werden hingegen insgesamt kritischer wahrgenommen (Grillitsch, 2010), weshalb eine Beschäftigung damit je nach Einstellungen eher zu bewussten Positionierungen und somit zu unterschiedlichem Beurteilungshandeln führen könnte. Dies ist von besonderer Bedeutung, da die Einstellungen von Lehrkräften relativ stabil sind (z.B. Kunter & Pohlmann, 2015). Weitere Untersuchungen der Einstellungen von Lehrkräften zu Bildungsstandards an größeren Stichproben wären jedoch wünschenswert, um aktuelle Entwicklungen dieser Einstellungen im Kontext der fortschreitenden Implementation abzubilden.

Die BNO wurde erstmals als Ergebnis einer Person-Situation-Interaktion untersucht. Es konnte gezeigt werden, dass die Einstellungen der Referendare zu Bildungsstandards (als Merkmal der Person) im Zusammenspiel mit der Konfrontation mit diesem Konzept (als situationaler Kontext) einen Einfluss auf die BNO haben. Je nachdem, ob die Standards abgelehnt oder befürwortet wurden, zogen Lehrkräfte im Vorbereitungsdienst bevorzugt die individuelle oder soziale Bezugsnorm zur Beurteilung heran. Dieser Befund schließt an frühere Studien an, die verdeutlichen, dass Reaktionen von Lehrkräften auf Reformen wie Bildungsstandards mit ihren Einstellungsprofilen zusammenhängen (Pant et al., 2008; Pöhlmann, Pant, Frenzel, Roppelt & Köller, 2014). Darüber hinaus ist dies relevant für den Diskurs zu den Auswirkungen des möglichen Spannungsverhältnisses von Inklusion und Standardorientierung (Teumer, 2012). Vorliegende Studien zum potentiellen Widerspruch

zwischen Inklusion und Standardorientierung bzw. Individualisierung und Standardisierung zeigten, dass Lehrkräfte nur teilweise eine Unvereinbarkeit zwischen beiden wahrnehmen (Kühn & Racherbäumer, 2013). Unsere Befunde weisen darauf hin, dass hierbei auch die Einstellungen zu berücksichtigen sind, da diese entweder zu einem mit Bildungsstandards übereinstimmenden oder aber diesem entgegengesetzten Lehrerhandeln führten. Wie Pöhlmann et al. (2014) zeigten, hat eine verstärkte Auseinandersetzung mit den Bildungsstandards u.a. den positiven Effekt, dass Lehrkräfte öfter kompetenzorientierte Schülertätigkeiten im Unterricht anregten. Dass die Auseinandersetzung mit Bildungsstandards bei Lehrenden mit einer positiven Einstellung zu diesen zu einer geringeren IBNO führt, deutet darauf hin, dass die Implementation von Standards und die Förderung einer diese befürwortenden Einstellung auch unintendierte „Nebenwirkungen“ haben könnte. Die IBNO wirkt sich schließlich, vermittelt über das damit einhergehende Lehrverhalten, günstig auf zahlreiche motivationale und affektive Schülermerkmale aus (Köller, 2005). Für inklusive Leistungsbeurteilung betont Prengel (2014) darüber hinaus, dass die Kombination unterschiedlicher BNO neu zu bedenken ist.

Weil Bildungsstandards eine am kriterialen Maßstab orientierte Leistungsbewertung implizieren, wurde in der vorliegenden Studie auch die selten untersuchte KBNO in den Blick genommen. Während die IBNO und die SBNO in der KBA gut abbildbar waren, fand sich die KBNO in ihrer reinen Form lediglich bei einer Person. Eine mögliche Erklärung dafür ist, dass Beurteilungen nach SBNO und KBNO nicht klar voneinander unterscheidbar sind. Eine Beurteilung nach der einen oder anderen Norm äußert sich in der KBA lediglich in unterschiedlichen Minimal- und Maximalbewertungen der besten und schlechtesten fiktiven Testergebnisse. Auch wurde die KBA nicht zur Messung der KBNO konzipiert und die Modifikation, dass minimal 0 und maximal 100 Punkte erreichbar sind, hob den kriterialen Maßstab eventuell nicht prominent genug hervor. Entsprechend könnte das seltene

Auftreten der KBNO in unserer Studie auch darauf zurückzuführen sein, dass diese nicht optimal operationalisiert wurde. Zukünftige Forschung würde von der Entwicklung anderer verhaltensnaher Operationalisierungen der KBNO sehr profitieren.

Dass die kriteriale Norm kaum angewendet wurde, spiegelt jedoch auch die Befunde vorliegender Forschung zur diagnostischen Kompetenz von Lehrkräften wider. In diesen Studien (z.B. Leucht, Tiffin-Richards, Vock, Pant & Köller, 2015) wurde gezeigt, dass Lehrkräfte vor allem ein klasseninternes Referenzsystem anwenden und ihre Urteile über verschiedene Klassen hinweg kaum vergleichbar sind; was darauf zurückgeführt wird, dass ihnen ein externer kriterialer Referenzrahmen zur Leistungsbewertung fehle (Leucht et al., 2015). Entsprechend könnte vermutet werden, dass nicht nur in der experimentellen Situation, sondern auch in der Unterrichtspraxis vor allem die individuelle und soziale, nicht aber die kriteriale Bezugsnorm eine Rolle spielen. Warum Lehrkräfte die kriteriale Bezugsnorm so selten anwenden, ist nicht geklärt. Möglicherweise widerspricht es dem Gerechtigkeitsgefühl, den besten Lernenden aus einer Gruppe die bestmögliche Beurteilung vorzuenthalten, nur weil sie nicht die maximal mögliche Leistung gezeigt haben. Hier deutet sich an, dass kriteriale Normen wie Bildungsstandards wahrscheinlicher bei der Leistungsbeurteilung einbezogen werden würden, wenn diese eine hinreichende Passung zur faktischen Leistungsstreuung in der zu beurteilenden Gruppe aufweisen.

Es ist zu berücksichtigen, dass sich unsere Befunde ausschließlich auf Berliner Lehrkräfte im Vorbereitungsdienst beziehen. Ob diese auf andere Stichproben, wie z.B. Lehrkräfte mit langjähriger Berufserfahrung, übertragbar sind, bleibt offen. Weiterhin ist anzumerken, dass die zur Messung der BNO herangezogene KBA eine relativ künstliche Bewertungssituation darstellt, die nur wenig mit der tatsächlichen Praxis der Leistungsbewertung übereinstimmt, in der Lehrkräfte i.d.R. über weitere Informationen über die Lernenden ver-

fügen. Diese Einschränkung der ökologischen Validität beinhaltet gleichzeitig eine Erhöhung der internen Validität, da mit der KBA die BNO relativ unkonfundiert erhoben werden kann (Wilbert & Gerdes, 2009). Weiterhin wird auf diese Weise das Beurteilungsverhalten erfasst, was gegenüber Selbstaussagen zur BNO weniger anfällig für Verzerrungen in sozial erwünschte Richtung ist.

Zusammengefasst liefert unsere Studie erste Ergebnisse zum potentiellen Spannungsverhältnis von Inklusion und Standardorientierung, welches in der Literatur diskutiert wird, aber erst in Ansätzen empirisch untersucht ist (z.B. Kühn & Racherbäumer, 2013). Zukünftige Forschung sollte dieses Spannungsverhältnis nicht nur in Bezug auf die Leistungsbeurteilung, sondern auch in Bezug auf weitere Aspekte des Unterrichtshandelns (z.B. der jeweiligen Unterrichtsgestaltung) untersuchen. Eine bewusste Reflexion der mit diesen Konstrukten eventuell assoziierten Widersprüche könnte den Lehrkräften ermöglichen, in der Unterrichtspraxis situationsspezifisch die Stärken beider Ansätze zu nutzen, ohne hierbei Konflikte mit den eigenen Überzeugungen und Einstellungen zu erleben.

## 7 Literatur

- Ainscow, M., Booth, T. & Dyson, A. (2006a). *Improving schools, developing inclusion*. London: Routledge.
- Ainscow, M., Booth, T. & Dyson, A. (2006b). Inclusion and the standards agenda. Negotiating policy pressures in England. *International Journal of Inclusive Education*, 10, 295–308.
- Altrichter, H. (2011). Bildungsstandards und Individualisierung im Unterricht – Zwei Reformpolitiken im Spannungsfeld. In F. Hofmann, D. Martinek & U. Schwantner (Hrsg.), *Binnendifferenzierter Unterricht und Bildungsstandards – (k)ein Widerspruch?* (S. 9–27). Wien: Lit.
- Avramidis, E. & Norwich, B. (2002). Teachers' attitudes towards integration / inclusion. A review of the literature. *European Journal of Special Needs Education*, 17, 129–147.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469–520.
- Beer, R. (2006). *Bildungsstandards. Einstellungen von Lehrerinnen und Lehrern*. Wien: Lit.
- Biewer, G. (2012). Die neue Welt der Bildungsstandards und ihre erziehungswissenschaftliche Rezeption aus der Perspektive einer Inklusiven Pädagogik. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 81, 9–21.
- Buschkühle, C.-P., Duncker, L. & Oswald, V. (Hrsg.). (2009). *Bildung zwischen Standardisierung und Heterogenität. Ein interdisziplinärer Diskurs*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dickhäuser, O. & Rheinberg, F. (2003). Bezugsnormorientierung: Erfassung, Probleme, Perspektiven. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Selbstkonzept und Motivation und Selbstregulation* (S. 41–55). Göttingen: Hogrefe.
- Emmrich, R., Ernst, C.-M., Harych, P. & Wesselhöft, K. (2012). *Vergleichsarbeiten der Jahrgangsstufe 8 in Berlin als Beitrag zur Schul- und Unterrichtsentwicklung*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Eversheim, U. (2015). Bildungsstandards versus Inklusion? Probleme und Potentiale der Standard- und Kompetenzorientierung für einen inklusiven (Sport)Unterricht. In S. Meier (Hrsg.), *Inklusion als Herausforderung. Aufgabe und Chance für den Schulsport* (S. 207–224). Berlin: Logos.
- Fazio, R. H., Chen, J., McDonel, E. C. & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, 18, 339–357.
- Freudenthaler, H. H. & Specht, W. (2005). *Bildungsstandards aus Sicht der Anwender. Evaluation der Pilotphase I zur Umsetzung nationaler Bildungsstandards in der Sekundarstufe I*. Graz: Zentrum für Schulentwicklung.
- Grillitsch, M. (2010). *Bildungsstandards auf dem Weg in die Praxis. Ergebnisse einer Befragung von Lehrkräften und Schulleiter/innen der Sekundarstufe I zur Rezeption der Bildungsstandards und deren Implementation*. Graz: Leykam.

- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.
- Heinrich, M. (2010). Zum Problem der Anerkennung fragiler Bildungsprozesse innerhalb neuer Steuerung und demokratischer Governance. In S. Aufenanger, F. Hamburger, R. Ludwit & R. Tippelt (Hrsg.), *Bildung in der Demokratie. Beiträge zum 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft* (S. 125–143). Opladen: Budrich.
- Helmke, A. & Hosenfeld, I. (2003). Vergleichsarbeiten (VERA): eine Standortbestimmung zur Sicherung schulischer Kompetenzen – Teil 1: Grundlagen, Ziele, Realisierung. *Schulverwaltung NRW, Ausgabe Hessen/Rheinland-Pfalz/Saarland* (1), 10–13.
- Hofmann, F. (2011). Bildungsstandards und Individualisierung – ein Widerspruch? Unterschiedliche Reaktionsmuster von Lehrpersonen – ein Hypothesen generierender Beitrag. In F. Hofmann, D. Martinek & U. Schwantner (Hrsg.), *Binnendifferenzierter Unterricht und Bildungsstandards – (k)ein Widerspruch?* (S. 47–62). Wien: Lit.
- Holmeier, M. (2012). Bezugsnormorientierung im Unterricht im Kontext zentraler Abiturprüfungen. In K. M. Merki (Hrsg.), *Zentralabitur* (S. 237–261). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Jantowski, A. (2014). Beurteilung von Schülerleistungen im Spannungsfeld von Standardorientierung und individueller Förderung. In A. Jantowski & R. Möllers (Hrsg.), *Unterricht im Spannungsfeld zwischen Kompetenz- und Standardorientierung* (S. 26–34). Bad Berka: Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien.
- Kessels, U., Erbring, S. & Heiermann, L. (2014). Implizite Einstellungen von Lehramtsstudierenden zur Inklusion. *Psychologie in Erziehung und Unterricht*, 61, 189–202.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters Kluwer.
- Köller, O. (2005). Bezugsnormorientierung von Lehrkräften: Konzeptuelle Grundlagen, empirische Befunde und Ratschläge für praktisches Handeln. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendungen* (S. 189–202). Stuttgart: Kohlhammer.
- Kühn, S. M. & Racherbäumer, K. (2013). Standardisierung und/oder Individualisierung? Empirische Befunde zur Umsetzung von Maßnahmen zur individuellen Förderung im Kontext zentraler Abschlussprüfungen. *Unterrichtswissenschaft*, 41, 173–190.



- Kullmann, H., Lütje-Klose, B., Textor, A., Berard, J. & Schitow, K. (2014). Inklusiver Unterricht – (Auch) eine Frage der Einstellung! Eine Interviewstudie über Einstellungen und Bereitschaften von Lehrkräften und Schulleitungen zur Inklusion. *Schulpädagogik heute*, 5 (10), 1–14.
- Kunter, M. & Pohlmann, B. (2015). Lehrer. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 261-282). Heidelberg: Springer.
- Leucht, M., Tiffin-Richards, S., Vock, M., Pant, H. A. & Köller, O. (2012). Diagnostische Kompetenz von Englischlehrkräften. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44, 163–177.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.
- Mischo, C. & Rheinberg, F. (1995). Erziehungsziele von Lehrern und individuelle Bezugsnormen der Leistungsbewertung. *Zeitschrift für Pädagogische Psychologie*, 9, 139–151.
- Oelkers, J. & Reusser, K. (2008). *Qualität entwickeln – Standards sichern – mit Differenz umgehen*. Bonn: Bundesministerium für Bildung und Forschung.
- Pant, H. A., Stanat, P., Pöhlmann, C., Hecht, M., Jansen, M., Kampa, N., Lenski, A. E., Penk, C., Radmann, S., Roppelt, A., Schroeders, U., Siegle, T., Ziemke, A. (2013). Der Blick in die Länder. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 159–247). Berlin: Waxmann.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 64, 827–845.
- Paulus, C. (2013). *Einstellungen zu Inklusion: Die deutsche Fassung des My-Thinking-About-Inklusion (MTAI) Fragebogens*. Zugriff am 10.02.2016 unter <http://scidok.sulb.uni-saarland.de/volltexte/2013/5554/pdf/MTAI.pdf>
- Pöhlmann, C., Pant, H. A., Frenzel, J., Roppelt, A. & Köller, O. (2014). Auswirkungen einer Intervention auf die Auseinandersetzung und Arbeit mit Bildungsstandards bei Mathematik-Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 17, 113–133.
- Prenzel, A. (2014). Inklusive Bildung: Grundlagen, Praxis, offene Fragen. In T. Häcker & M. Whalm (Hrsg.), *Inklusion in Schule und Lehrer\_innenbildung* (S. 1–19). Bad Heilbrunn: Klinkhardt.
- Racherbäumer, K. & Kühn, S. M. (2013). Zentrale Prüfungen und individuelle Förderung. *Zeitschrift für Bildungsforschung*, 3, 27–45.
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Göttingen: Hogrefe.
- Rheinberg, F. (1998). Bezugsnorm-Orientierung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 39–43). Weinheim: Beltz.
- Richard, F. D., Bond, C. F. & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, 7, 331–363.
- Sacher, W. (2009). *Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe*. Bad Heilbrunn: Klinkhardt.

- Sander, A. (2005). Bildungsstandards und Bildungsbarrieren: Thesen aus Perspektive einer inklusiven Pädagogik. In U. Geiling & A. Hinz (Hrsg.), *Integrationspädagogik im Diskurs. Auf dem Weg zu einer inklusiven Pädagogik?* (S. 110–113). Bad Heilbrunn: Klinkhardt.
- Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2004). Zielorientierung und Bezugsnormorientierung. Zum Zusammenhang zweier Konzepte. *Zeitschrift für Pädagogische Psychologie*, 18, 93–99.
- Seitz, S. (2006). Inklusive Didaktik nach PISA. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 75, 192–199.
- Stoiber, K. C., Gettinger, M. & Goetz, D. (1998). Exploring factors influencing parents' and early childhood practitioners' beliefs about inclusion. *Early Childhood Research Quarterly*, 13, 107–124.
- Teumer, S. (2012). Paradoxe Handlungsanforderungen an LehrerInnen im Spannungsfeld der Neugestaltung des Schulanfangs. *Zeitschrift für Inklusion* (3). Zugriff am 10.02.2016 unter <http://www.inklusion-online.net/index.php/inklusion-online/article/view/47/47>
- Trumpa, S., Janz, F., Heyl, V. & Seifried, S. (2014). Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. *Zeitschrift für Bildungsforschung*, 4, 241–256.
- Wilbert, J. & Gerdes, H. (2009). Die Bezugsnormwahl bei der Bewertung schulischer Leistungen durch angehende Lehrkräfte des Förderschwerpunktes Lernen. *Heilpädagogische Forschung*, 35, 122–132.
- Wilhelm, M. (2009). *Integration in der Sekundarstufe I und II. Wie die Umsetzung im Fachunterricht gelingt*. Weinheim, Basel: Beltz.

### **4.3 Studie 3: Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective**

Holder, K. & Kessels, U. (2017). Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective. *Social Psychology of Education*, 20, 471–490.

**Die finale Publikation ist verfügbar bei Springer unter/**

**The final publication is available at [link.springer.com](https://doi.org/10.1007/s11218-017-9384-z) via**

**<https://doi.org/10.1007/s11218-017-9384-z>**

### Abstract

Teacher judgments of student performance might be biased by stereotypes, which can result in disadvantages for members of negatively stereotyped social groups. On the basis of the shifting standards theory, we examined gender and ethnic biases in student teachers' judgments. According to this theory, whether such judgment biases are masked or revealed depends on the judgment context. Specifically, the use of subjective scales allows different standards of judgment to be applied to members of different social groups and should therefore mask biases, whereas the use of objective scales does not allow for a standard shift and should therefore reveal stereotyped expectations. The present studies tested predictions derived from this theory by varying the judgment context: using "educational standards" as a setting that requires objective scales and "inclusion" as a setting that employs subjective scales. Study 1 explored the stereotype that girls are not as good in mathematics as boys, Study 2 the stereotype that students of Turkish origin are not as good in German as nonimmigrants. Participants in two experimental vignette studies (Study 1: N = 155; Study 2: N = 265) were asked to evaluate the performance of a stereotyped or nonstereotyped target student in a context requiring ratings on either objective or subjective scales (2 x 2 between-subjects design). In line with the shifting standards theory, stereotyped students were evaluated similarly to nonstereotyped students on subjective scales but were evaluated more negatively on objective scales. Results indicate that student teachers' judgments are biased by gender and ethnic stereotypes.

**Keywords:** Shifting standards, Gender stereotypes, Ethnic stereotypes, Teacher judgments

## 5 Gesamtdiskussion

Ziel der vorliegenden Arbeit ist es, Erkenntnisse zu gewinnen, wie sich die in der Literatur diskutierten potentiellen Gegensätze zwischen Inklusion und Bildungsstandards aus der Sicht von Lehrkräften darstellen. Dabei wurden Aspekte der Unterrichtsgestaltung ebenso untersucht wie Aspekte der Leistungsbeurteilung und damit zusammenhängende Konsequenzen, die mit der Hervorhebung unterschiedlicher Beurteilungsmaßstäbe im Kontext beider Reformen für stereotypisierte Schülergruppen verbunden sein können. Basierend auf dem theoretischen und empirischen Erkenntnisstand zu Implikationen für Unterrichtsgestaltung und Leistungsbeurteilung, die mit Inklusion und Bildungsstandards verbunden sein können, und dem bildungswissenschaftlichen Diskurs um die Vereinbarkeit beider Reformen, wurden folgende drei Fragestellungen entwickelt und in empirischen Studien überprüft:

1. Verbinden Lehrkräfte mit Inklusion und Bildungsstandards unterschiedliche Implikationen für Unterrichtsgestaltung und Leistungsbeurteilung?
2. Führt die Auseinandersetzung mit dem Thema Bildungsstandards oder mit dem Thema Inklusion bei Lehrkräften je nach Einstellung zu diesen Reformen zur Anwendung unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung?
3. Welche Konsequenzen gehen mit der Anwendung subjektiver und objektiver Beurteilungsmaßstäbe für die Leistungsbeurteilung stereotypisierter Schülergruppen einher?

Nachfolgend werden zunächst die zentralen Ergebnisse der drei Studien in Bezug auf die Beantwortung dieser Fragestellungen zusammengefasst. Anschließend erfolgt eine Einordnung und Reflexion der Ergebnisse vor dem Hintergrund bisheriger Forschung und im Kontext aktueller bildungswissenschaftlicher Diskurslinien, bevor deren Limitationen

beleuchtet werden. Weiterhin werden aus den Ergebnissen Implikationen für die pädagogische Praxis und für weitere Forschung abgeleitet. Ein studienübergreifendes Fazit bildet schließlich den Abschluss der Arbeit.

## **5.1 Zusammenfassung zentraler Ergebnisse**

Die experimentell angelegte Studie 1 („*Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven und standardorientierten Unterricht aus der Sicht von Lehrkräften*“) untersuchte, ob und inwiefern Lehrkräfte mit Inklusion und Bildungsstandards eine jeweils andere Unterrichtsgestaltung und Leistungsbeurteilung verbinden. Unterrichtsgestaltung wurde durch zentrale Merkmale auf Ebene der Sicht- und Tiefenstrukturen von Unterricht (Oser & Baeriswyl, 2001) operationalisiert und Leistungsbeurteilung durch die drei Formen der BNO (Rheinberg, 1980). Dabei wurden für die Implementation beider Reformen relevante Lehrermerkmale als Kontrollvariablen und zwei unterschiedliche Beschreibungsperspektiven von Unterricht (Selbst- und Fremdeinschätzung) als zusätzliche Einflussfaktoren berücksichtigt. Insgesamt zeigte sich, dass sich aus der Sicht von Lehrkräften inklusiver Unterricht von standardorientiertem Unterricht hinsichtlich der Umsetzung zentraler Merkmale von Unterrichtsqualität und hinsichtlich der darin auftretenden BNO bei der Leistungsbeurteilung unterscheidet. Berichtete Unterschiede waren jedoch zum Teil auf Einstellungen zu und Erfahrungen mit beiden Reformen sowie diesbezügliche Selbstwirksamkeitserwartungen zurückzuführen und variierten mit der Perspektive, aus der Unterricht und Leistungsbeurteilung beschrieben wurden.

Betrachtet man nur die Beschreibung des Unterrichts einer *anderen* (fiktiven) Lehrkraft, die wahrscheinlich den jeweiligen „Prototypen“ vom Unterrichten in beiden Kontexten widerspiegelt, zeigten sich unter Kontrolle der genannten lehrerseitigen Merkmale, dass mit inklusivem Unterricht ein höheres Maß an Differenzierung, eine stärker

ausgeprägte Unterstützung durch die Lehrkraft und eine stärker ausgeprägte individuelle BNO sowie eine geringer ausgeprägte kriteriale BNO und tendenziell eine weniger effiziente Klassenführung verbunden wurden als mit standardorientiertem Unterricht. Somit konnte erstmals empirisch gezeigt werden, dass Lehrkräfte auf theoretisch-konzeptioneller Ebene mit Inklusion und Bildungsstandards unterschiedliche Implikationen für Unterrichtsgestaltung und Leistungsbeurteilung verbinden, wengleich angegebene Unterschiede zwischen beiden Reformansätzen scheinbar teilweise auf lehrerseitige Merkmale zurückgehen. Ungeklärt blieb dabei die Frage, aus welchem Grund sich in der Beschreibung des eigenen Unterrichts weniger Unterschiede zwischen Inklusion und Standardorientierung zeigten als in der Beschreibung des Unterrichts einer anderen Lehrkraft. Vermutet wurde jedoch, dass Einschätzungen des eigenen Unterrichts in sozial erwünschte oder selbstwertdienliche Richtung verzerrt sein könnten (z.B. Wubbels et al., 1992). Somit könnten Unterschiede, die Lehrkräfte mit beiden Reformen verbinden, durch die Beschreibung „guten Unterrichts“ unabhängig von beiden Reformen überdeckt werden (siehe Diskussionskapitel zu Studie 1, S. 102).

Die ebenfalls experimentell angelegte Studie 2 (*„Lehrkräfte zwischen Bildungsstandards und Inklusion: Eine experimentelle Studie zum Einfluss von „Standardisierung“ und „Individualisierung“ auf die Bezugsnormorientierung“*) untersuchte mögliche Auswirkungen von Inklusion und Bildungsstandards auf die Leistungsbeurteilung. Es wurde geprüft, inwiefern die Auseinandersetzung mit Inklusion oder Bildungsstandards in Abhängigkeit von den Einstellungen zu diesen Reformen zur Anwendung unterschiedlicher Bezugsnormen (Rheinberg, 1980) führt. Die Ergebnisse zeigten, dass eine Auseinandersetzung mit Bildungsstandards bei Lehrkräften mit einer positiven Einstellung zu diesen eine schwächer ausgeprägte individuelle BNO sowie tendenziell eine stärker ausgeprägte soziale BNO nach sich zieht. Bei Personen mit einer

negativen Einstellung zu Bildungsstandards führte die Auseinandersetzung mit diesem Reformansatz zu einer stärker ausgeprägten individuellen BNO und tendenziell zu einer schwächer ausgeprägten sozialen BNO. Erwartungswidrig konnte für die Auseinandersetzung mit Inklusion je nach Einstellungen zu Inklusion kein Einfluss auf die BNO nachgewiesen werden. Dass die Auseinandersetzung mit Bildungsstandards bei positiven Einstellungen zu diesen zu einer geringeren Ausprägung der individuellen BNO und einer tendenziell stärkeren Ausprägung der sozialen BNO führte, kann jedoch den Implikationen von Inklusion für die Leistungsbeurteilung konträr gegenüberstehen. Da lediglich eine Person die kriteriale BNO bei der Leistungsbeurteilung anwendete, konnten die darauf bezogenen Teilhypothesen in Studie 2 nicht überprüft werden.

Da im Kontext beider Reformen unterschiedliche Maßstäbe für die Leistungsbeurteilung hervorgehoben werden, untersuchte die dritte Studie („*Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective*“) mögliche Konsequenzen der Anwendung subjektiver und objektiver Beurteilungsmaßstäbe für die Leistungsbeurteilung stereotypisierter Schülergruppen. Die Studie gliederte sich in zwei Teilstudien. Es wurde jeweils überprüft, ob die Anwendung einer subjektiven Skala (im Kontext eines individualisierten Beurteilungssettings) und die Anwendung einer objektiven Skala (im Kontext eines standardisierten Beurteilungssettings) zu unterschiedlichen Beurteilungen von stereotypisierten gegenüber nicht stereotypisierten Lernenden führt. Die Ergebnisse beider Teilstudien zeigten, dass Lernende, die einer stereotypisierten Gruppe angehören (Mädchen in Mathematik und türkischstämmige Schüler in Deutsch) gegenüber der jeweils nicht stereotypisierten Schülergruppe (Jungen in Mathematik und deutschstämmige Schüler in Deutsch) für die identische gezeigte Leistung negativer beurteilt wurden, wenn die Einschätzung anhand der objektiven Messskala der Bildungsstandards erfolgte. Bei Verwendung einer subjektiven Skala im Kontext der



Einschätzung eines Portfolios, zeigten sich hingegen keine Unterschiede in der Leistungsbeurteilung zwischen Mitgliedern stereotypisierter und nicht stereotypisierter Schülergruppen. Somit scheinen bei der Leistungsbeurteilung stereotypenbasierte *standard shifts* stattzufinden. Die Ergebnisse deuten darauf hin, dass das ggf. mit den Bildungsstandards verbundene zunehmende Denken in objektiven Beurteilungsmaßstäben dazu führen kann, dass die Leistungen stereotypisierter Schülerinnen und Schüler unzutreffend beurteilt werden, auch wenn dies zunächst kontraintuitiv erscheint. Anzumerken ist jedoch, dass die mit Inklusion verbundene Denkweise in individuell subjektiven Beurteilungsmaßstäben stereotypengeleitete Beurteilungen ebenfalls nicht verringert. Das Auftreten solcher Beurteilungen ist durch die Anwendung subjektiver Skalen lediglich nach außen hin, z.B. für Schülerinnen und Schüler, weniger sichtbar.

### **5.2 Einordnung und Reflexion der Ergebnisse**

Die Ergebnisse der drei Studien haben Erkenntnisse dazu beigetragen, inwieweit Lehrkräfte Inklusion und Bildungsstandards in Bezug auf Unterrichtsgestaltung und Leistungsbeurteilung als vereinbar ansehen und welche Konsequenzen, mit der Hervorhebung unterschiedlicher Beurteilungsmaßstäbe im Kontext beider Reformen für die Leistungsbeurteilung stereotypisierter Schülergruppen einhergehen können. Weiterhin zeigen die Ergebnisse auf, dass lehrerseitige Merkmale in Bezug auf die wahrgenommene Vereinbarkeit und die gleichzeitige Umsetzung beider Reformen zu berücksichtigen sind. In den folgenden Abschnitten erfolgt eine Einordnung und Reflexion der zentralen Ergebnisse, die deren Interpretation in Bezug auf weitere Forschungsergebnisse und aktuell identifizierte Diskurslinien einschließt.

### **5.2.1 Die Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Unterrichtsgestaltung aus der Sicht von Lehrkräften**

Die Ergebnisse aus Studie 1 weisen insgesamt darauf hin, dass Lehrkräfte auch unter Kontrolle relevanter lehrerseitiger Merkmale mit Inklusion und Bildungsstandards eine jeweils unterschiedliche Ausgestaltung bestimmter Unterrichtsmerkmale auf Ebene der Sicht- und Tiefenstrukturen verbinden. Berichtete Unterschiede scheinen jedoch teilweise auf die Einstellungen zu und das Ausmaß an Erfahrungen mit beiden Reformen sowie darauf bezogene Selbstwirksamkeitserwartungen zurückzugehen. Daher kann angenommen werden, dass je nach Ausprägung dieser lehrerseitigen Merkmale die Unterrichtsgestaltung im Kontext beider Reformen mehr oder weniger unterschiedlich eingeschätzt wird. Es könnte beispielsweise vermutet werden, dass hohe Selbstwirksamkeitserwartungen bezogen auf die Gestaltung inklusiven Unterrichts bei gleichzeitig hohen Selbstwirksamkeitserwartungen bezogen auf die Gestaltung standardorientierten Unterrichts mit weniger berichteten Unterschieden zwischen beiden Reformen einhergehen.

Wird den untersuchten unterschiedlichen Beurteilungsperspektiven auf Unterrichtsgestaltung Rechnung getragen, zeigte sich zudem, dass Unterschiede zwischen beiden Reformansätzen vorrangig auf theoretisch-konzeptioneller Ebene und scheinbar weniger auf der Ebene des eigenen Unterrichtshandelns wahrgenommen werden. Da mögliche Erklärungen für letzteren Befund bereits im Diskussionskapitel zu Studie 1 dargestellt wurden (siehe S. 102) werden im Folgenden jene von den Lehrkräften berichteten Unterschiede zwischen inklusivem und standardorientiertem Unterricht diskutiert, die sich unter Kontrolle der genannten lehrerseitigen Merkmale auf theoretisch-konzeptioneller Ebene zeigten.

Hinsichtlich der Umsetzung von Methodenvielfalt, Frontalunterricht und kognitiver Aktivierung nahmen Lehrkräfte unter Kontrolle der erhobenen lehrerseitigen Faktoren auf

theoretisch-konzeptioneller Ebene keine Unterschiede zwischen inklusivem und standardorientierten Unterricht wahr. Dies deutet darauf hin, dass beide Reformen aus der Sicht von Lehrkräften einerseits gleichermaßen die Realisierung unterschiedlicher Methoden zulassen und andererseits das Anregen tiefgehender Prozesse des Nachdenkens und der aktiven Auseinandersetzung mit einem Lerngegenstand sowohl im inklusiven als auch im standardorientierten Unterricht möglich erscheint. Für die Ausgestaltung dieser Merkmale von Unterrichtsqualität scheinen folglich aus Sicht der Lehrkräfte keine unterschiedlichen Implikationen mit beiden Reformen einherzugehen. Im Diskurs um die Vereinbarkeit von Inklusion und Bildungsstandards wird von verschiedenen Autorinnen und Autoren die Befürchtung geäußert, mit der Einführung der Bildungsstandards könnte eine Verengung des Curriculums oder die Verbreitung von *teaching to the test* einhergehen (Biewer, 2012; Moser Opitz, 2006; Seitz & Platte, 2006). Da vermutet werden kann, dass ein methodisch abwechslungsreicher und kognitiv aktivierender Unterricht vor dem Hintergrund einer solchen Engführung des Curriculums oder im Rahmen von ausgeprägtem *teaching to the test* nur bedingt umsetzbar wäre, scheinen sich diese Befürchtungen nicht in den Angaben der befragten Lehrkräfte widerzuspiegeln. Dies wird auch durch Ergebnisse früherer Studien zur Rezeption der VERA unterstützt, die verdeutlichten, dass aus Sicht von Lehrkräften keine Verengung des Curriculums mit diesen einhergeht (Richter et al., 2014; Wacker & Kramer, 2012).

Ein weiterer wichtiger Befund bezüglich des Diskurses um die Vereinbarkeit von Inklusion und Bildungsstandards ist, dass unter Kontrolle der genannten lehrerseitigen Merkmale auf theoretisch-konzeptioneller Ebene mit standardorientiertem gegenüber inklusivem Unterricht eine schwächer ausgeprägte Unterstützung durch die Lehrkraft verbunden wurde. Dieser Befund könnte damit in Verbindung stehen, dass die befragten Lehrkräfte die in der Literatur diskutierte Annahme teilen, eine Orientierung an einheitlichen Lernzielen

im Unterricht könnte zur Vernachlässigung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf führen (Biewer, 2012; Eversheim, 2015; Moser Opitz, 2006; Ritter & Hennies, 2015; Seitz, 2006b). Jedoch lag auch der Mittelwert der mit standardorientiertem Unterricht verbundenen Unterstützung deutlich über dem theoretischen Mittelwert der zur Erfassung eingesetzten Skala (siehe deskriptive Ergebnisse aus Studie 1, S. 87). Dies deutet darauf hin, dass aus Sicht der befragten Lehrkräfte auch eine Orientierung an Bildungsstandards im Unterricht mit einer stark ausgeprägten Unterstützung durch die Lehrkraft einhergeht, die auch aus der Perspektive von Inklusion bedeutsam scheint.

Darüber hinaus weisen die Ergebnisse aus Studie 1 darauf hin, dass Lehrkräfte unter Kontrolle der genannten lehrerseitigen Merkmale auf theoretisch-konzeptioneller Ebene die Klassenführung im inklusiven Unterricht tendenziell weniger effizient einschätzen als im standardorientierten Unterricht. Im dargestellten bildungswissenschaftlichen Diskurs um die Vereinbarkeit beider Reformen wird dieses Merkmal der Unterrichtsgestaltung nicht gesondert betrachtet. Dass Lehrkräfte die Klassenführung im inklusiven Unterricht vergleichsweise wenig effizient einschätzen, ist dennoch von Relevanz, da Merkmalen effizienter Klassenführung eine wichtige Bedeutung für inklusive Lerngruppen zukommt (z.B. Emmer & Stough, 2001). Wie frühere Studien aufzeigen, stehen Lehrkräfte insbesondere der Inklusion von Schülerinnen und Schülern mit Verhaltensauffälligkeiten, z.B. im Förderschwerpunkt „emotionale und soziale Entwicklung“, kritischer gegenüber als beispielsweise der Inklusion von Lernenden mit körperlich-motorischen Einschränkungen (Avramidis & Norwich, 2002; Eichfeld & Algermissen, 2016; Schwab & Seifert, 2015). Da, wie Hillenbrand (2011, S. 26) beschreibt, Verhaltensstörungen aus didaktischer Perspektive Unterrichtsstörungen darstellen, könnten die befragten Lehrkräfte eine höhere Anzahl Unterrichtsstörungen und daher eine weniger effiziente Klassenführung mit inklusivem Unter-

richt verbinden. Dieser Befund zeigte sich in Studie 1 zwar auch unter Kontrolle der Einstellung zu Inklusion, jedoch wurde zur Erhebung dieser Einstellung eine relativ allgemeine Skala eingesetzt, die sich nicht auf Einstellungen zur Inklusion bestimmter Schülergruppen bezieht (siehe eingesetzte Erhebungsinstrumente in Studie 1, S. 88). Somit könnten nicht erfasste Einstellungen, z.B. gegenüber der Inklusion verhaltensauffälliger Lernender, diesen Befund erklären. Offen bleibt jedoch, ob Lehrkräfte auch tatsächlich eine wenig effiziente Klassenführung im inklusiven Unterricht umsetzen. Denn die Ergebnisse von Henke, Koch und Spörer (2015) zeigten z.B. auf, dass Schülerinnen und Schülern an inklusiven Pilot-schulen die im Unterricht realisierte Klassenführung als gut einschätzten.

Schließlich ist für den Diskurs um potentielle Spannungsverhältnisse zwischen beiden Reformen insbesondere der Befund bedeutsam, dass auch unter Kontrolle der relevanten Lehrermerkmale auf theoretisch-konzeptioneller Ebene mit standardorientiertem Unterricht eine geringer ausgeprägte innere Differenzierung verbunden wurde als mit inklusivem Unterricht. Dies verweist auf die aus unterschiedlichen Perspektiven beleuchtete Frage, inwiefern im Rahmen der Vorgabe von Bildungsstandards eine für inklusiven Unterricht angemessene innere Differenzierung umsetzbar ist (z.B. Seitz, 2006b; Werning & Arndt, 2015). Dass Lehrkräfte mit einer Orientierung an den Bildungsstandards weniger Differenzierung im Unterricht verbinden als mit Inklusion, deutet darauf hin, dass mit der Vorgabe einheitlicher Lernziele eine stärkere Gleichförmigkeit des Unterrichts assoziiert wird, wenn gleich diese nicht intendiert ist (z.B. KMK, 2004a). Anknüpfend an den Diskurs um die Vereinbarkeit beider Reformen könnte dies einerseits der von Inklusionspädagoginnen und Inklusionspädagogen diskutierten zu geringen Differenzierung der Kompetenzstufenmodelle geschuldet sein, die sich nicht als Grundlage eignen der Heterogenität im inklusiven Unterricht angemessen zu begegnen (Seitz, 2006b; Seitz & Platte, 2006). Andererseits

könnte dies auch auf die in der Literatur geäußerte Vermutung zurückgehen, dass Bildungsstandards durch die Vorgabe einheitlicher Lernziele grundsätzlich dem Gedanken des ziel-differenten Unterrichtens entgegenstehen können (Werning & Arndt, 2015).

Dennoch lag, ähnlich wie bei der Unterstützung durch die Lehrkraft, der Mittelwert der mit standardorientiertem Unterricht verbundenen Differenzierung über dem theoretischen Mittelwert der zu Erfassung eingesetzten Skala (siehe deskriptive Ergebnisse aus Studie 1, S. 86). Zudem fiel die Effektstärke des gefundenen Unterschieds zwischen der jeweils mit Inklusion und Bildungsstandards verbundenen Differenzierung im Unterricht relativ gering aus. Folglich bleibt offen, ob und inwiefern dieser von den Lehrkräften beschriebene Unterschied tatsächlich eine Unvereinbarkeit darstellt. Nichtsdestotrotz ist dieser Befund anschlussfähig an Hinweise aus früheren Studien, dass Lehrkräfte eine differenzierende Unterrichtsgestaltung im Kontext von Standardisierungsprozessen als eingeschränkt erleben (Beer, 2006; Racherbäumer & Kühn, 2013).

Dass Lehrkräfte standardorientiertem Unterricht eine geringer ausgeprägte Differenzierung beimessen als inklusivem Unterricht, lässt zunächst keine Schlüsse über die tatsächliche Umsetzung dieser im Kontext der Gleichzeitigkeit beider Reformen zu. Es liegen zwar erste Hinweise vor, dass u.U. sowohl im inklusiven als auch im standardorientierten Unterricht eine differenzierende Unterrichtsgestaltung realisiert wird (Liebers et al., 2014; Richter et al., 2014), jedoch zeigten verschiedene Studien, dass Lehrkräfte in Deutschland im Allgemeinen kaum binnendifferenzierende Maßnahmen im Unterricht umsetzen (Lankes & Carstensen, 2007; Schiepe-Tiska et al., 2016), und zwar selbst dann nicht, wenn sie in besonders leistungsheterogenen Klassen unterrichten (Warwas, Hertel & Labuhn, 2011). Dennoch wurde ebenfalls herausgestellt, dass Lehrkräfte individuelle Förderung mehrheitlich als erstrebenswert ansehen und dieser einen hohen Stellenwert beimessen (Solzbacher, 2008). Folglich scheinen Lehrkräfte sich zwar der

Notwendigkeit innerer Differenzierung bewusst zu sein, diese jedoch aktuell im Unterricht kaum umzusetzen.

Als mögliche Gründe dafür werden u.a. hinderliche schulische Rahmenbedingungen diskutiert (Trautmann & Wischer, 2011; Vock & Gronostaj, 2017). Trautmann und Wischer (2011) weisen z.B. darauf hin, dass im Unterricht nicht nur die optimale Förderung aller Lernenden zu berücksichtigen ist, sondern auch Anforderungen, die sich aus gesellschaftlichen Funktionen der Schule ergeben, wobei aus Sicht der Autorinnen und Autoren vor allem die Selektions- und Allokationsfunktion für die Umsetzung differenzierender Maßnahmen erschwerend erscheint. Durch die Steigerung der Heterogenität der Schülerschaft um Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf deutet sich eine Verschärfung dieser Problematik an, wie von Inklusionspädagoginnen und Inklusionspädagogen angedeutet (z.B. Büker & Meier, 2017; Seifert & Müller-Zastrau, 2014). Instrumente zur Erfüllung der Selektions- und Allokationsfunktion von Schule sind beispielsweise Abschlussprüfungen, die auf den KMK-Bildungsstandards basieren (z.B. KMK, 2004g, KMK, 2012b). Auch Zielkonflikte, die durch die gleichzeitig erforderliche „Orientierung an gleichen Standards für alle“ (Trautmann & Wischer, 2011, S. 127) und an individuellen Lernvoraussetzungen entstehen, werden als Herausforderung für die Umsetzung differenzierender Maßnahmen im Unterricht und als mögliche Ursache für deren geringes Ausmaß angeführt.

Folglich könnten Bildungsstandards sowohl vermittelt über darauf basierenden Abschlussprüfungen als auch aufgrund ihrer Orientierungsfunktion für die Gestaltung von Unterrichtsprozessen (Klieme et al., 2003) einschränkend für die Umsetzung differenzierender Maßnahmen im Unterricht wirken, wenngleich dies nicht intendiert ist. Die Ergebnisse aus Studie 1 scheinen dies im Hinblick auf inklusiven Unterricht zu bestätigen. Dieser Problematik könnte eventuell durch eine Erhöhung der Passung der in

den Bildungsstandards formulierten Kompetenzanforderungen an die heterogenen Lernvoraussetzungen und -möglichkeiten der Schülerschaft in inklusiven Klassen Rechnung getragen werden.

Zusammenfassend kann hinsichtlich der Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Unterrichtsgestaltung festgehalten werden, dass Lehrkräfte mit beiden Reformen auch unter Kontrolle lehrerseitiger Faktoren auf theoretisch-konzeptioneller Ebene eine jeweils unterschiedliche Ausgestaltung bestimmter Merkmale auf Ebene der Sicht- und Tiefenstrukturen von Unterricht assoziieren. Jedoch bleibt offen, inwiefern die von den Lehrkräften beschriebenen Unterschiede sich tatsächlich als Unvereinbarkeiten darstellen, da Lehrkräfte mit beiden Reformen, gemessen an den theoretischen Mittelwerten der eingesetzten Skalen, eine relativ hohe Unterrichtsqualität zu verbinden scheinen. Die verschiedenen theoretischen Ansätze inklusiven und standardorientierten Unterrichts deuten ebenfalls darauf hin, dass mit *beiden* Reformen Verbesserungen der Unterrichtsqualität einhergehen sollen. Mit Blick auf dieses gemeinsame übergeordnete Ziel und vor dem Hintergrund der gefundenen Ergebnisse könnte das stärkere Herstellen von Bezügen zwischen beiden Reformstrategien in Bezug auf Unterrichtsgestaltung gewinnbringend sein.

### **5.2.2 Die Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Leistungsbeurteilung aus der Sicht von Lehrkräften**

Hinsichtlich der Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Leistungsbeurteilung deuten die Ergebnisse aus Studie 1 zunächst darauf hin, dass aus der Sicht von Lehrkräften mit beiden Reformen jeweils unterschiedlich starke Ausprägungen der individuellen und kriterialen BNO einhergehen. Unter Kontrolle der Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards und der diesbezüglichen Selbstwirksamkeitserwartungen verbanden Lehrkräfte mit inklusivem Unterricht auf theoretisch-konzeptioneller Ebene eine stärker ausgeprägte individuelle BNO und eine schwächer ausgeprägte



kriteriale BNO als mit standardorientiertem Unterricht. Standardorientiertem Unterricht attestierten sie umgekehrt eine schwächer ausgeprägte individuelle BNO und eine stärker ausgeprägte kriteriale BNO als inklusivem Unterricht. Folglich scheinen Lehrkräfte mit beiden Reformansätzen vorrangig die jeweils auch theoretisch primär erscheinenden Bezugsnormen (z.B. Klieme et al., 2003; Seifert & Müller-Zastrau, 2014) zu verbinden, die einander tendenziell konträr gegenüberstehen (Holmeier, 2013).

Wie die Ergebnisse aus Studie 2 aufzeigen, schlagen sich diese assoziierten Unterschiede nicht genau deckungsgleich in der tatsächlichen Anwendung dieser Bezugsnormen in Abhängigkeit von den Einstellungen zu Inklusion und Bildungsstandards nieder, denn vor allem die kriteriale BNO wurde von den Teilnehmenden kaum angewendet. Mögliche Gründe dafür wurden bereits im Diskussionskapitel zu Studie 2 (siehe S. 132f.) angeführt. Allerdings zeigte Studie 2, dass die Auseinandersetzung mit Bildungsstandards bei positiven Einstellungen zu diesen zu einer geringeren Ausprägung der individuellen BNO und tendenziell zu einer stärkeren Ausprägung der sozialen BNO führt. Je nach Einstellungen zu den Bildungsstandards kann das damit zusammenhängende Beurteilungshandeln folglich einer primären Orientierung an der individuellen Bezugsnorm, wie im Kontext von Inklusion angedacht (z.B. Seitz, 2006b), widersprechen. Die tendenziell stärkere Ausprägung der sozialen BNO nach Auseinandersetzung mit den Bildungsstandards und einer positiven Einstellung zu diesen steht nicht nur den Leitgedanken inklusiver Leistungsbeurteilung, sondern auch der geäußerten Intention entgegen, mit den Standards und VERA die Perspektive von Lehrkräften über soziale, klasseninterne Vergleiche hinaus stärker auf eine kriteriale Bezugsnorm auszurichten (z.B. Oelkers & Reusser, 2008). Ergebnisse früherer Studien zeigten, dass positive Einstellungen zu den Standards und zu VERA eine wichtige Rolle für deren gelingende Umsetzung spielen können (z.B. Richter et al., 2014) und deshalb gefördert werden sollten. Jedoch kann, wie die Ergebnisse aus Studie 2 andeuten, das Fördern

einer positiven Einstellung zu Bildungsstandards aus der Perspektive inklusiver Leistungsbeurteilung auch weniger erwünschte Folgen haben, wenn dadurch die Orientierung an der individuellen Bezugsnorm vernachlässigt und die Orientierung an der sozialen Bezugsnorm verstärkt wird.

Folglich deuten die Ergebnisse aus Studie 1 und Studie 2 darauf hin, dass aus der Sicht von Lehrkräften mit den Reformansätzen Inklusion und Bildungsstandards potentiell gegensätzliche Anforderungen an die Leistungsbeurteilung verbunden sein können. Somit spiegeln die Ergebnisse das in der Literatur beschriebene Spannungsverhältnis zwischen der Wertschätzung individueller Lernprozesse einerseits und der Orientierung an einheitlichen, übergreifenden Leistungsnormen (bzw. sozialen Vergleichsgruppen) andererseits (Amrhein, 2014; Bräu, 2008; Teumer, 2012) wider, dass durch die Gleichzeitigkeit beider Reformen besonders deutlich hervortritt. Dies erscheint auch deshalb von Bedeutung, da das Erleben gegensätzlicher oder mehrdeutiger Anforderungen (*ambiguity*) u.a. als mögliche Ursache für Stress- und Belastungserleben im Lehrerberuf angesehen wird (Rothland, 2013; Wiley, 2000).

Wie gravierend dieses wahrgenommene Spannungsverhältnis zwischen Inklusion und Bildungsstandards ausfällt, hängt jedoch mit der Absolutheit der Forderung einer individuellen BNO bei der Leistungsbeurteilung im inklusiven Unterricht zusammen. Während einige Autorinnen und Autoren die stärkere oder *vorrangige* Orientierung an der individuellen Bezugsnorm bei der Leistungsbeurteilung postulieren (z.B. Seifert & Müller-Zastrau, 2014), ist aus der Sicht anderer mit Inklusion ein elementarer Wandel im Schulsystem bzw. in der Gesellschaft verbunden, innerhalb dessen weitestgehend auf Kategorisierungen, Etikettierungen und Vergleiche von Schülerinnen und Schülern verzichtet werden soll (z.B. Hinz, 2009; Wocken, 2012). Aus dieser Perspektive würde *einzig* das Anwenden einer individuellen BNO notwendig und legitim erscheinen. Wie

Ahrbeck (2015) konstatiert, könnten jedoch Vorbehalte gegenüber Bildungsstandards und eine ausschließliche Anwendung der individuellen Bezugsnorm den gesellschaftlichen Bezug der Schule schwächen, da Leistungen dadurch lediglich aus einem eingeschränkten Blickwinkel betrachtet würden. Auch vor dem Hintergrund, dass Lehrkräfte aufgrund der Allokations- und Selektionsfunktion von Schule trotz der pädagogischen Sinnhaftigkeit der individuellen BNO letztlich stets vergleichend bewerten müssen (Katzenbach, 2010; Vock & Gronostaj, 2017), erscheint das ausschließliche Anwenden der individuellen Bezugsnorm auch im inklusiven Unterricht weder möglich noch sinnvoll, solange diese Funktion von Schule besteht. Dennoch ist Inklusion untrennbar damit verbunden, der Individualität von Lernprozessen bzw. individuellen Voraussetzungen und Möglichkeiten der Schülerinnen und Schüler auch bei der Leistungsbeurteilung Rechnung zu tragen. Folglich scheint nicht nur aufgrund der Gleichzeitigkeit von Inklusion und Bildungsstandards, sondern auch aufgrund der Selektions- und Allokationsfunktion von Schule das in Studie 1 und Studie 2 identifizierte Spannungsverhältnis zwischen der Orientierung an individuellen Lernverläufen und -voraussetzungen einerseits und an übergreifenden kriterialen und/oder sozialen Bezugsnormen andererseits kaum auflösbar zu sein.

Darüber hinaus kann vermutet werden, dass aus der Perspektive der Schülerinnen und Schüler alle drei Bezugsnormen ihre Berechtigung haben können. Das Anlegen einer individuellen Bezugsnorm hängt insbesondere positiv mit motivationalen und affektiven Schülermerkmalen zusammen (für einen Überblick siehe Köller, 2005; Mischo & Rheinberg, 1995). Damit geht jedoch auch die Gefahr einher, dass diese ihr Können falsch einschätzen, da kein externer Maßstab vorliegt, mit dem die eigenen Fähigkeiten verglichen werden können (Gläser-Zikuda, 2010; Holmeier, 2013). Die kriteriale Bezugsnorm basiert auf einem solchen Maßstab und ermöglicht in Bezug auf bestimmte Lernziele eine realisti-

sche Einschätzung der eigenen Leistungen. Wie Sacher (2009) argumentiert, kann das Anlegen der kriterialen Bezugsnorm aber auch zu Überforderungsgefühlen führen und die Motivation beeinträchtigen, wenn die angelegten Kriterien z.B. für leistungsschwächere Schülerinnen und Schüler unerfüllbar erscheinen. Trotz der negativen Auswirkungen der sozialen BNO (z.B. Dickhäuser et al., 2017) kann vor dem Hintergrund der Theorie sozialer Vergleiche (Festinger, 1954) angenommen werden, dass Schülerinnen und Schüler auch das Bedürfnis haben, sich selbst bzw. ihre Leistungen mit anderen zu vergleichen. Dass Schülerinnen und Schüler auch tatsächlich soziale Vergleiche ziehen, belegen z.B. zahlreiche Studien zum Big Fish Little Pond Effect (BFLPE; Marsh, 1994), der den Einfluss sozialer Vergleiche auf das akademische Selbstkonzept beschreibt (z.B. Lüdtke, Köller, Marsh & Trautwein, 2005; Marsh & Hau, 2003). In einigen Studien wurde jedoch nachgewiesen, dass soziale Vergleiche im Sinne des BFLPE für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Regelklassen mit problematischen Folgen einhergehen können. Darin zeigte sich, dass das Begabungskonzept von Schülerinnen und Schülern mit Förderbedarf aufgrund sozialer Vergleiche und damit zusammenhängender Bezugsgruppeneffekte in Regelklassen niedriger ausfiel als in Sonderklassen (Haeberlin, 1991; Möller, 2013) und dass ein negativer Bezugsgruppeneffekt auf das akademische Selbstkonzept von Kindern mit Förderbedarf in Regelklassen auftrat, während sich dieser in Förderklassen nicht zeigte (Nusser & Wolter, 2016). Zudem ist anzumerken, dass Schülerinnen und Schüler auf Basis der sozialen Bezugsnorm je nach sozialer Vergleichsgruppe unterschiedliche Einschätzungen und keine Informationen über die individuelle Lernentwicklung oder das Erreichen bestimmter Lernziele erhalten (Gläser-Zikuda, 2010; Holmeier, 2013; Sacher, 2009).

Sowohl vor dem Hintergrund der verschiedenen Funktionen von Schule als auch aus der Perspektive von Schülerinnen und Schülern erscheint folglich eine Kombination unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung zur Abmilderung des in Studie 1

und Studie 2 aufgezeigten Spannungsverhältnisses zwischen Inklusion und Bildungsstandards sinnvoll. Eine solche Kombination verschiedener Bezugsnormen wird auch unabhängig von der Gleichzeitigkeit dieser beiden Reformen als gewinnbringend betrachtet (Blömeke, Herzig & Tulodziecki, 2007; Gläser-Zikuda, 2010; Klieme, 2003). Da die drei Bezugsnormen einander jedoch tendenziell konträr gegenüberstehen, wird darauf hingewiesen, dass eine Umsetzung einer solchen Kombination durchaus mit Schwierigkeiten verbunden sein kann (Holmeier, 2013).

Zur Ausbalancierung dieser potentiell gegensätzlichen Anforderungen an die Leistungsbeurteilung, könnte insbesondere das Etablieren eines mehrperspektivischen Leistungsbegriffs, wie von Prenzel (2015a, 2015b) vorgeschlagen, aussichtsreich sein. Die Autorin beschreibt, dass speziell bei der inklusiven Leistungsbeurteilung drei Bezugsnormen kombiniert werden sollten. Zunächst sollte mit einer „universellen Bezugsnorm“ (Prenzel, 2015b, S. 53) grundlegend eine leistungsunabhängige Anerkennung aller Lernenden als Teil der Lerngemeinschaft sichergestellt werden. Darüber hinaus erscheint auch das Anlegen einer „kriterial-sozial-vergleichenden Bezugsnorm“ (Prenzel, 2015a, S. 53) nach Ansicht der Autorin notwendig. Diese soll Transparenz darüber schaffen, wie das eigene Lernergebnis im Vergleich zu den Lernergebnissen Anderer zu verorten ist und welche der vorgegebenen Leistungsnormen zu einem bestimmten Zeitpunkt erreicht wurden (Prenzel, 2015a). Denn auch in inklusiven Klassen sollte nach Prenzel (2015b) das Vorhandensein unterschiedlicher Leistungsstände nicht tabuisiert, jedoch von diskriminierenden Zuschreibungen gelöst werden. Schließlich sieht Prenzel (2015a, S. 53) es als zentral an, eine „kriterial-individuelle“ Bezugsnorm anzuwenden. Damit ist gemeint, keine hierarchisierenden Beurteilungen vorzunehmen, sondern jedes Kind als „auf seiner Stufe kompetent“ (Prenzel, 2015b, S. 41) anzuerkennen. Diese kriterial-individuelle Bezugsnorm soll sich auf kriteriale

Kompetenzanforderungen beziehen, die fachspezifisch das gesamte mögliche Leistungsspektrum umfassen (Prengel, 2015b). Mit einem solchen Leistungsverständnis könnte folglich eine Zusammenführung der potentiell widersprüchlichen Bezugsnormen einhergehen, die Lehrkräfte, wie Studie 1 und Studie 2 zeigten, jeweils bei der Leistungsbeurteilung im Kontext von Inklusion und Bildungsstandards wahrnehmen und anwenden. Es kann zudem vermutet werden, dass sich auch die Kompetenzstufenmodelle der Bildungsstandards als kriteriale Bezugsrahmen für inklusive Leistungsbeurteilung eignen könnten, jedoch nur unter der Prämisse, dass diese sich adäquat zur differenzierten Abbildung sehr breiter Leistungsspektren eignen.

In Studie 3 wurde die Unterschiedlichkeit zwischen im Rahmen von Inklusion und Bildungsstandards hervorgehobenen Beurteilungsmaßstäben zur Untersuchung stereotypenbasierter Leistungsbeurteilung genutzt. Damit konnte, auch über den Diskurs um die Vereinbarkeit beider Reformen hinaus, der in Deutschland relativ begrenzte Forschungsstand zu vorliegenden negativen Stereotypen von Lehrkräften gegenüber Mädchen in Mathematik (z.B. Nürnberger et al., 2016; Tiedemann, 2002) und türkischstämmigen Schülern in Deutsch (z.B. Glock & Krolak-Schwerdt, 2013; Lorenz et al., 2016) und deren Einfluss auf die Leistungsbeurteilung erweitert werden.

Bezogen auf die Vereinbarkeit von Inklusion und Bildungsstandards deuten die Ergebnisse aus Studie 3 darauf hin, dass ein Denken in objektiven Beurteilungsmaßstäben, auch wenn es zunächst kontraintuitiv erscheint, mit einer ungerechteren Leistungsbeurteilung von Angehörigen stereotypisierter Schülergruppen einhergehen kann. Denn durch das Einnehmen einer übergreifenden Beurteilungsperspektive werden die Leistungen stereotypisierter Lernender mit den Leistungen *aller* anderen Lernenden verglichen und eine objektiv gute Leistung erscheint relativ unwahrscheinlicher. Dies äußert sich, wie Studie 3 hervorbrachte, in einer negativeren Beurteilung von stereotypisierten

gegenüber nicht stereotypisierten Lernenden auf objektiven Skalen, selbst wenn diese die identische Leistung zeigen. Eine solche Benachteiligung in der Leistungsbeurteilung aufgrund von Stereotypen steht grundlegend der inklusiven Idee der Anerkennung aller Schülerinnen und Schüler in ihrer Individualität und der Wertschätzung von Unterschiedlichkeit konträr gegenüber (z.B. Carle, 2017; Kullmann et al., 2014; Werning, 2014). Werning (2011, S. 5) beschreibt gar die „Überbetonung einer verobjektivierten Leistungsmessung“ als eines der Hauptprobleme bei der Umsetzung schulischer Inklusion. Weiterhin erscheint dieser Befund bedeutsam, da einzelne Studien darauf hindeuten, dass bei Lehrkräften auch negative Stereotype über Lernende mit Lern- oder Leistungsbeeinträchtigungen vorzuliegen scheinen (z.B. Cornett-Ruiz & Hendricks, 1993; Foster, Schmidt & Sabatino, 1976). Auch Stereotype gegenüber diesen Schülergruppen könnten folglich durch das Anlegen objektiver Maßstäbe bei der Leistungsbeurteilung im Sinne der Shifting-Standards-Theorie (Biernat et al., 1991; Biernat, 2009) sichtbar werden, wenngleich bisher eher selten untersucht wurde, inwiefern lehrerseitige Stereotype gegenüber Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf vorhanden sind.

Zudem muss beachtet werden, dass die mit Inklusion verbundene Denkweise in individuell subjektiven Beurteilungsmaßstäben das Auftreten stereotypengeleiteter Beurteilungen ebenfalls nicht reduziert, sondern diese lediglich nach außen hin weniger sichtbar werden. Denn die Anwendung einer subjektiven Skala bei der Leistungsbeurteilung im Sinne der Shifting-Standards-Theorie ist von der Anwendung der bereits diskutierten individuellen Bezugsnorm abzugrenzen. Wie die Shifting-Standards-Theorie verdeutlicht, ruft die Beurteilung auf einer subjektiven, nicht kriterial verankerten Skala nicht die Fokussierung auf die zu beurteilende Person und deren Voraussetzungen und Möglichkeiten hervor, sondern eine Orientierung an den Leistungserwartungen an die soziale Gruppe, welcher

diese Person angehört (z.B. Biernat & Manis, 1994). Folglich geht eine Beurteilung auf subjektiven Skalen mit dem Anlegen einer sozialen Bezugsnorm einher. So meint die Einschätzung „sehr gut“ bei einer stereotypisierten Schülerin oder einem stereotypisierten Schüler, dass ihre oder seine Leistung lediglich im Vergleich zu anderen Mitgliedern ihrer oder seiner sozialen Gruppe als sehr gut angesehen wird (im Sinne von „für eine/n wie dich sehr gut“) und hat lediglich eine begrenzte Aussagekraft. Wenn Schülerinnen und Schüler unterschiedlichen, ggf. stereotypisierten sozialen Gruppen angehören, stellt sich zudem die Frage, welche soziale Gruppe den Vergleichshorizont bei der Leistungsbeurteilung bildet. Denkbar ist beispielsweise, dass eine Schülerin mit sonderpädagogischem Förderbedarf bei der Beurteilung auf subjektiven Skalen lediglich mit anderen Schülerinnen mit (dem gleichen) sonderpädagogischen Förderbedarf verglichen wird. Eine ausschließlich auf subjektiven Skalen verankerte Leistungsbeurteilung im Kontext von Inklusion kann folglich ähnlich wie eine ausschließliche Orientierung an der individuellen Bezugsnorm problematisch erscheinen. Da, wie dargestellt wurde, auch das Anlegen objektiver Beurteilungsmaßstäbe mit negativen Konsequenzen für die Leistungsbeurteilung stereotypisierter Lernender einhergehen kann, scheinen die Ergebnisse aus Studie 3 ebenfalls zu bekräftigen, dass unterschiedliche Maßstäbe bei der Leistungsbeurteilung und auch bei der Leistungsrückmeldung kombiniert werden sollten.

Zusammenfassend verweisen die gefundenen Ergebnisse darauf, dass mit Inklusion und Bildungsstandards in Bezug auf Leistungsbeurteilung ein Spannungsverhältnis verbunden ist, das Lehrkräfte möglicherweise vor gegensätzliche Anforderungen stellt. Zudem zeigen die Ergebnisse auf, welche Folgen das Einnehmen subjektiver und objektiver Perspektiven bei der Leistungsbeurteilung, die jeweils durch Inklusion und Bildungsstandards hervorgehoben werden, für stereotypisierte Schülergruppen haben kann. Eine vollständige Auflösung dieses mit Inklusion und Bildungsstandards verbundenen



Spannungsverhältnisses erscheint jedoch kaum realistisch. Hinweise auf die Sinnhaftigkeit der Kombination verschiedener Bezugsnormen bei der Leistungsbeurteilung lassen jedoch die von den Lehrkräften wahrgenommenen Gegensätze zwischen Inklusion und Bildungsstandards auch als Chance erscheinen, hinsichtlich der Leistungsbeurteilung voneinander zu profitieren. Sowohl im Diskurs um die Vereinbarkeit von Inklusion und Bildungsstandards als auch in Bezug auf die gleichzeitige Umsetzung beider Reformen könnte es daher aussichtsreich sein, sich weniger auf potentielle Gegensätze, sondern eher auf die Kombination der Stärken beider Reformansätze in Bezug auf Leistungsbeurteilung zu konzentrieren.

### **5.3 Limitationen**

Die drei durchgeführten Studien stellen wertvolle Erkenntnisse zur Vereinbarkeit von Inklusion und Bildungsstandards aus der Sicht von Lehrkräften bereit. Die gefundenen und diskutierten Ergebnisse unterliegen jedoch verschiedenen Limitationen, die deren Belastbarkeit und Generalisierbarkeit einschränken. Da in den Diskussionskapiteln der einzelnen Studien bereits die jeweils spezifischen Einschränkungen herausgearbeitet wurden (Studie 1: S. 107; Studie 2: S. 137f.; Studie 3: S. 167), werden nachfolgend noch einmal zentrale, studienübergreifende Limitationen kritisch reflektiert. Diese beziehen sich vor allem auf die verwendeten Untersuchungsdesigns und die untersuchten Stichproben.

Zu berücksichtigen ist zunächst, dass alle drei durchgeführten Studien experimentell angelegt sind. Experimentelle Studien weisen eine relativ hohe interne Validität auf und ermöglichen die Untersuchung kausaler Wirkmechanismen. Mit einer hohen internen Validität gehen jedoch Einschränkungen der externen Validität einher. Wenngleich alle drei Studien in Feld- und nicht in Laborumgebungen durchgeführt wurden, wurden darin künstliche Situationen geschaffen, die nur bedingt auf die schulische Praxis übertragbar sind.

Zum einen erfolgte eine Gegenüberstellung von Inklusion und Bildungsstandards als voneinander getrennte Unterrichtskontexte (Studie 1 und Studie 2). Diese sind jedoch in der Realität nicht trennscharf auseinanderzuhalten, da deren Einflüsse miteinander vermischt und zeitgleich auftreten. Zum anderen sollten die Leistungen fiktiver Schülerinnen und Schüler beurteilt werden (Studie 2 und Studie 3), ohne dass den Befragten weitere Informationen über diese Lernenden vorlagen. In der schulischen Praxis stehen Lehrkräften im Regelfall jedoch mehrere Informationen über die zu beurteilenden Lernenden zur Verfügung, die in die Leistungsbeurteilung einfließen können. Um zu überprüfen, wie ökologisch valide und verlässlich die gefundenen Ergebnisse wirklich sind, erscheint es folglich sinnvoll, entsprechende Feldstudien durchzuführen, die sich auf die tatsächliche Unterrichtsgestaltung und Leistungsbeurteilung im Kontext von Inklusion und Bildungsstandards beziehen. Insbesondere die Ergebnisse aus Studie 3 sind auf ihre externe Validität zu überprüfen, da unzutreffende, auf Stereotypen basierende Lehrererwartungen in der Realität durch den Kontakt mit den betreffenden Schülerinnen und Schülern akkurater werden sollten (Raudenbush, 1984).

Weiterhin muss die Einschränkung angeführt werden, dass allen drei Studien Gelegenheitsstichproben zugrunde liegen und keine Zufallsstichproben gezogen wurden. Dieser Limitation könnte mit Untersuchungen der Fragestellungen dieser Arbeit an großen, repräsentativen Stichproben begegnet werden. Praktisch wäre dies jedoch nur durch Sekundäranalysen bereits vorliegender repräsentativer Datensätze oder das Erheben entsprechender Daten innerhalb größerer Forschungsverbände umsetzbar. Bereits vorliegende Datensätze enthalten die für die Fragestellungen dieser Arbeit relevanten Konstrukte jedoch (wenn überhaupt) nur in Teilen. So untersucht z.B. das Nationale Bildungspanel (NEPS) Einstellungen von Lehrkräften zu Inklusion und Merkmale der Unterrichtsgestaltung, jedoch nicht die Einstellungen zu Bildungsstandards.

Damit geht zudem die Limitation einher, dass sich die gefundenen Ergebnisse ausschließlich auf Lehrkräfte im Vorbereitungsdienst (Studie 1 und Studie 2) bzw. Lehramtsstudierende (Studie 3) aus Berlin und Brandenburg beziehen. Ob und inwiefern diese auf andere Stichproben, wie z.B. Lehrkräfte mit langjähriger Berufserfahrung oder Lehrkräfte aus anderen Bundesländern, übertragbar sind, ist eine offene Frage. Aufgrund dessen erscheint es sinnvoll, diese Fragestellungen anhand weiterer, unterschiedlicher Stichproben zu untersuchen, um die Verlässlichkeit und Generalisierbarkeit der gefundenen Ergebnisse zu untermauern. Dies ist insbesondere deshalb von Bedeutung, da sich verschiedene Bundesländer sowohl in der Art als auch im Umfang der Implementation von Bildungsstandards und vor allem von Inklusion unterscheiden (siehe Kapitel 1). Auch diesen unterschiedlichen Gegebenheiten müsste in weiteren Untersuchungen Rechnung getragen werden, um die Generalisierbarkeit entsprechender Ergebnisse zu erhöhen.

### **5.4 Implikationen**

Die Ergebnisse der drei Studien weisen auf unterschiedliche Aspekte hin, hinsichtlich derer aus der Sicht von Lehrkräften möglicherweise eine eingeschränkte Vereinbarkeit von Inklusion und Bildungsstandards gegeben ist und zeigen auf, welche Folgen das Einnehmen unterschiedlicher Perspektiven bei der Leistungsbeurteilung für stereotypisierte Schülergruppen haben kann. Weiterhin deuten die Ergebnisse darauf hin, dass bestimmte Lehrermerkmale in Bezug auf die wahrgenommene Vereinbarkeit und die gleichzeitige Umsetzung beider Reformen eine Rolle zu spielen scheinen. Aus den dargestellten und diskutierten Ergebnissen werden nachfolgend zunächst Vorschläge für die pädagogische Praxis abgeleitet, die sich auf Lehrkräfte und die Lehrerbildung beziehen. Zudem ergeben sich, wie teilweise bereits im vorigen Abschnitt angedeutet, einige Ansatzpunkte für die weitere Forschung, die ebenfalls nachfolgend herausgearbeitet werden.

### 5.4.1 Praxis

Für die pädagogische Praxis erscheint zunächst der Befund aus Studie 1 bedeutsam, dass mit beiden Reformen aus Lehrersicht zwar teilweise unterschiedliche Anforderungen an Unterrichtsgestaltung verbunden sein können, jedoch gleichzeitig auch, gemessen an den theoretischen Mittelwerten der eingesetzten Skalen, eine relativ hohe Unterrichtsqualität. Da die verschiedenen theoretisch-konzeptionellen Ansätze inklusiven und standardorientierten Unterrichts ebenfalls den Schluss zulassen, dass mit *beiden* Reformen Verbesserungen der Unterrichtsqualität intendiert sind (z.B. Drieschner, 2009; Kullmann et al., 2014), erscheint es lohnend, dass Lehrkräfte sich differenziert mit diesen auseinandersetzen. Dies könnte z.B. bereits in der Lehramtsausbildung realisiert werden. Ein besseres Verständnis der intendierten Wirkungen der Bildungsstandards und VERA für Unterrichtsentwicklung könnte auch für inklusiven Unterricht, z.B. hinsichtlich Möglichkeiten der Umsetzung kognitiver Aktivierung, gewinnbringend sein, wenngleich die Passung der Bildungsstandards zur tatsächlichen Heterogenität der Schülerinnen und Schüler möglicherweise eingeschränkt erscheint. Indem die jeweiligen Potenziale beider Reformansätze zur Herstellung einer hohen Unterrichtsqualität konstruktiv genutzt und miteinander in Verbindung gebracht werden, könnte ein Unterricht realisiert werden, der für *alle* Schülerinnen und Schüler positive Erträge hat (Seidel & Shavelson, 2007; Werning & Baumert, 2013).

Hinsichtlich der Leistungsbeurteilung weisen die Ergebnisse aus Studie 1 und Studie 2 darauf hin, dass sich Inklusion und Bildungsstandards konträr gegenüberstehen können und die Gleichzeitigkeit beider Reformen für Lehrkräfte mit auszubalancierenden Spannungsverhältnissen bezüglich der Anwendung von Bezugsnormen einhergehen kann. Ein kritisches Hinterfragen der assoziierten Gegensätze könnte den Lehrkräften ermöglichen, in der Beurteilungspraxis je nach Situation die Stärken beider Ansätze zu nutzen, ohne dabei in Konflikte mit den eigenen Einstellungen (Studie 2) zu geraten. Eine

Auseinandersetzung mit den Stärken beider Reformansätze könnte auch dazu beitragen, eine sinnvolle Kombination unterschiedlicher Bezugsnormen bei der Leistungsbeurteilung und auch bei der Leistungsrückmeldung an Schülerinnen und Schüler zu realisieren, damit diese Aufschluss über individuelle Lernfortschritte erhalten und gleichzeitig zu realistischen Selbsteinschätzungen kommen. Voraussetzung dafür ist jedoch eine Sensibilisierung von Lehrkräften für die spezifischen Vor- und Nachteile der mit Inklusion und Bildungsstandards implizierten Bezugsnormen, welche bereits in der Lehramtsausbildung erfolgen sollte. Bezüglich der Kombination verschiedener Bezugsnormen im inklusiven Unterricht können beispielsweise inklusionsgeeignete Kompetenzraster herangezogen werden (Schuck, 2014), wie sie z.B. innerhalb des diagnostischen Instruments „Individuelle Lernstandsanalysen“ (ILeA, z.B. Schnitzler, Scheerer-Neumann & Ritter, 2008) zur Verfügung gestellt werden. Dieses Instrument gibt Aufschluss über individuell erreichte Lernstände in Bezug auf kriterial festgelegte Kompetenzen, bezogen auf die individuelle Lernentwicklung und auch in Relation zu anderen Lernenden (Schnitzler et al., 2008).

Darüber hinaus deuten die Ergebnisse aus Studie 3 darauf hin, dass bei Lehrkräften negative Stereotype gegenüber Mädchen in Mathematik und gegenüber türkischstämmigen Schülern in Deutsch vorzuliegen scheinen. Daraus kann für die schulische Praxis der Vorschlag abgeleitet werden, dass Lehrkräfte eine erhöhte Aufmerksamkeit auf möglicherweise vorliegende stereotype Vorstellungen legen sollten, um sich des Einflusses dieser auf die Leistungsbeurteilung bewusst zu werden. Lorenz et al. (2016) sehen den Abbau von Stereotypen bei Lehrkräften als Strategie an, systematische Urteilsverzerrungen zu verringern. Zu einem solchen Abbau kann auch die Lehrkräfteaus- und -fortbildung einen Beitrag leisten, indem diese über möglicherweise vorliegende Stereotype und deren Entstehung informiert (Glock & Krolak-Schwerdt, 2013). Beispielsweise können gezielte Fortbildungen

dazu beitragen, dass Lehrkräfte ihre stereotypen Vorstellungen überdenken und zu kontrollieren lernen (Kawakami, Dovidio, Moll, Hermsen & Russin, 2000). Im Kontext der Gleichzeitigkeit von Inklusion und Bildungsstandards und dem Einnehmen damit verbundener Beurteilungsperspektiven erscheint es darüber hinaus ertragreich, dass Lehrkräfte sich der möglichen Auswirkungen verschiedener Beurteilungsformate auf die Urteilsbildung bewusstwerden.

### **5.4.2 Forschung**

Für die zukünftige Forschung ergeben sich Implikationen aus den gefundenen und diskutierten Ergebnissen, die sich einerseits auf die weitere Untersuchung potentiell gegensätzlicher Aspekte von Inklusion und Bildungsstandards richten, andererseits aber auch auf mögliche Beiträge der Forschung zu einer besseren wahrgenommenen Vereinbarkeit beider Reformen verweisen.

Studie 1 lieferte Erkenntnisse, hinsichtlich welcher Aspekte Lehrkräfte eine jeweils unterschiedliche Ausgestaltung von Unterricht und Leistungsbeurteilung mit Inklusion und Bildungsstandards verbinden. Da diesbezügliche Fragestellungen bisher kaum untersucht wurden, erscheint es sinnvoll weitere Studien anzuschließen, die die in der Literatur angenommene, potentielle Gegensätzlichkeit zwischen beiden Reformen aus der Sicht betroffener Akteure untersuchen. Dabei sollte auch eventuell unterschiedlichen Sichtweisen verschiedener Subgruppen von Lehrkräften Rechnung getragen werden, wie z.B. erfahrenen und weniger erfahrenen Lehrkräften, Lehrkräften unterschiedlicher Fächer und Schularten oder Lehrkräften mit und ohne sonderpädagogische Ausbildung. Denn beispielsweise zeigte sich in der Untersuchung von Einstellungen zu Inklusion, dass sich diese zwischen Studierenden und Berufsanfängern unterscheiden können (Hecht, Niedermair & Feyerer,

2016) und auch zwischen Lehrkräften an Sonderschulen und Grundschulen bzw. weiterführenden Schulen (Trumpa, Janz, Heyl & Seifried, 2014).

Da die Frage offenblieb, inwiefern berichtete Unterschiede zwischen Inklusion und Bildungsstandards tatsächlich auch in der Praxis wirksam werden, sollte zukünftige Forschung die tatsächliche Unterrichtsgestaltung und Leistungsbeurteilung im Kontext der Gleichzeitigkeit beider Reformen zum Gegenstand machen. Vor dem Hintergrund der Ergebnisse aus Studie 1 wäre beispielsweise eine genauere Betrachtung der Umsetzung innerer Differenzierung bzw. individueller Förderung im inklusiven Unterricht in Verbindung mit von Lehrkräften geäußerten Herausforderungen bei der Orientierung an Bildungsstandards lohnenswert. Ein Rückgriff auf methodisch mehrperspektivische Designs, die sowohl das Unterrichtshandeln als auch Selbstauskünfte von Lehrkräften erfassen, könnte dabei sinnvoll sein.

Studie 1 und Studie 2 berücksichtigten bei der Untersuchung der wahrgenommenen Vereinbarkeit von Inklusion und Bildungsstandards in Bezug auf Unterrichtsgestaltung und Leistungsbeurteilung auch jeweils die Einflüsse relevanter lehrerseitiger Merkmale. In der explorativ angelegten ersten Studie fungierten die Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie die Selbstwirksamkeitserwartungen bezogen auf inklusiven und standardorientierten Unterricht als Kontrollvariablen, wobei sich zeigte, dass berichtete Unterschiede zwischen beiden Reformansätzen scheinbar teilweise durch diese erklärt werden. Aus diesem Grund erscheint eine genauere Überprüfung der potentiellen Effekte und Wirkmechanismen dieser Faktoren in Bezug auf die wahrgenommene Vereinbarkeit beider Reformen lohnend sowie auch in Bezug auf deren gleichzeitige Umsetzung.

In Studie 2 wurde den Einstellungen zu Inklusion und Bildungsstandards in Bezug auf das Beurteilungsverhalten eine moderierende Rolle zugeschrieben. Die Ergebnisse

weisen einerseits darauf hin, dass mit einer positiven Einstellung zu den Bildungsstandards auch nicht intendierte „Nebenwirkungen“ einhergehen können, da eine Auseinandersetzung mit Bildungsstandards in Verbindung mit positiven Einstellungen zu diesen eine geringere individuelle und eine stärkere soziale BNO bei der Leistungsbeurteilung nach sich zog. Bisherige Studien bringen positive Einstellungen zu den Standards und VERA jedoch vorrangig mit deren gelingender Rezeption und Umsetzung in Verbindung (z.B. Richter et al., 2014). Die Ergebnisse aus Studie 2 könnten daher zum Anlass genommen werden auch nicht intendierte Nebeneffekte solch positiver Einstellungen, vor allem im Kontext der zunehmenden Umsetzung von Inklusion, zu untersuchen. Die Ergebnisse aus Studie 2 deuten andererseits darauf hin, dass die Einstellungen zu Inklusion keinen Einfluss auf die Anwendung unterschiedlicher Bezugsnormen nach der Auseinandersetzung mit dieser Reform zu haben scheinen, wobei mögliche Gründe bereits im Diskussionskapitel zu dieser Studie (siehe S. 131) beleuchtet wurden. Da die Wirkungen der Einstellungen von Lehrkräften zu Inklusion jedoch generell bisher kaum untersucht wurden (Gasterstädt & Urban, 2016; Sermier Dessemontet et al., 2011), sollten zukünftige Studien diese stärker fokussieren (de Boer et al., 2011). Dadurch könnten Erkenntnisse gewonnen werden, ob und wenn ja auf welche Weise positive Einstellungen und damit zusammenhängend auch hohe Selbstwirksamkeitserwartungen und vorliegende Erfahrungen die inklusive Unterrichtsgestaltung und Leistungsbeurteilung beeinflussen sowie darüber vermittelt eventuell auch kognitive, affektive und motivationale Schülermerkmale.

Die gefundenen Ergebnisse hinsichtlich des potentiell unterschiedlichen Ausmaßes einer differenzierenden Unterrichtsgestaltung im inklusiven gegenüber standardorientiertem Unterricht (Studie 1) und auch hinsichtlich wahrgenommener Spannungsverhältnisse zwischen beiden Reformen in Bezug auf Leistungsbeurteilung (Studie 1 und Studie 2) wurden ausführlich diskutiert. Es wurde herausgearbeitet, dass eine



Erhöhung der Passung von in den Bildungsstandards beschriebenen Kompetenzanforderungen zu den heterogenen Lernvoraussetzungen und -möglichkeiten der Schülerinnen und Schüler dazu beitragen könnte, dass sich Bildungsstandards stärker als Orientierungspunkte für die Planung und Durchführung inklusiven Unterrichts und auch als kriterialer Referenzrahmen inklusiver Leistungsbeurteilung eignen. Daraus lässt sich für zukünftige Forschung der Vorschlag ableiten, die Passung dieser Kompetenzanforderungen zu den Fähigkeiten von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf genauer zu untersuchen. Dies lässt sich z.B. durch die Einbeziehung jener Schülerinnen und Schüler in standardbasierte Tests realisieren, die ohnehin im Kontext der Herstellung einer uneingeschränkten Teilhabe an allgemeiner Bildung erforderlich scheint (Davies & Elliott, 2012; Gresch, Rjosk, Kocaj & Stanat, 2017). Erste diesbezügliche Forschungsergebnisse weisen darauf hin, dass bildungsstandardbezogene Testaufgaben zwar auch die Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf valide und reliabel erfassen konnten, jedoch eine zu geringe Differenzierungsfähigkeit im Spektrum relevanter Kompetenzen aufwiesen und eher zu schwierig ausfielen (Kocaj et al., 2016). Eine tendenziell zu hohe Schwierigkeit der Testaufgaben konnte auch in den Kompetenztestungen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf im NEPS nachgewiesen werden (Südkamp, Pohl, Hardt, Jordan & Duchhardt, 2015). Im Kontext der bildungsstandardbasierten Kompetenzerhebung gelang jedoch im IQB-Bildungstrend im Jahr 2016 die Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, die das gesamte Kompetenzspektrum dieser Lernenden abbildeten und zwischen verschiedenen Kompetenzständen zufriedenstellend differenzieren konnten (Böhme, Kohrt, Haag & Weirich, 2017). Weitere Evidenz zur Passung der in den Bildungsstandards beschriebenen Kompetenzanforderungen und darauf basierender Testaufgaben zu den

Lernvoraussetzungen und -möglichkeiten von Schülerinnen und Schülern mit unterschiedlichen sonderpädagogischen Förderbedarfen zu schaffen, erscheint äußerst relevant. Denn wenn die Kompetenzen jener Schülerinnen und Schüler in standardbasierten Kompetenzstufenmodellen adäquat abgebildet werden könnten, könnten diese auch für inklusiven Unterricht eine höhere Relevanz aufweisen, was auch zu einer Erhöhung der (wahrgenommenen) Vereinbarkeit beider Reformen beitragen könnte.

Da erste Befunde darauf hindeuten, dass bei Lehrkräften auch negative Stereotype über Lernende mit Lern- oder Leistungsbeeinträchtigungen vorzuliegen scheinen (Cornett-Ruiz & Hendricks, 1993; Foster, Schmidt & Sabatino, 1976) und in Studie 3 auf Basis der Shifting-Standards-Theorie (Biernat et al., 1991; Biernat, 2009) gezeigt wurde, dass Stereotype die Leistungsbeurteilung beeinflussen können, erscheint es außerdem lohnenswert, lehrerseitige Stereotype gegenüber Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in zukünftigen Studien näher zu untersuchen. Die nachgewiesenen negativere Einstellung von Lehrkräften zur Inklusion von Schülerinnen und Schülern mit Verhaltensauffälligkeiten (Avramidis & Norwich, 2002; Eichfeld & Algermissen, 2016; Schwab & Seifert, 2015), könnte ein Indiz dafür sein, dass vor allem über diese Schülergruppe negative Stereotype bzgl. ihrer Schulleistungen vorliegen können.

## **5.5 Fazit**

Ziel der vorliegenden Arbeit war es, Erkenntnisse zu gewinnen, inwiefern sich die in der Literatur diskutierte potentielle Gegensätzlichkeit zwischen Inklusion und Bildungsstandards aus der Sicht von Lehrkräften empirisch bestätigen lässt und welche Konsequenzen sich durch die Hervorhebung jeweils unterschiedlicher Beurteilungsmaßstäbe im Kontext dieser Reformen für die Leistungsbeurteilung stereotypisierter Schülergruppen ergeben. Unter Rückgriff auf verschiedene erziehungswissenschaftliche und sozialpsychologische

Theorien wurden bedeutsame Aspekte aufgezeigt, hinsichtlich derer aus der Sicht von Lehrkräften potentiell gegensätzliche Anforderungen mit beiden Reformen verbunden sein können.

In Bezug auf die *Unterrichtsgestaltung* scheinen Lehrkräfte auf theoretisch-konzeptioneller Ebene ein jeweils unterschiedliches Ausmaß innerer Differenzierung, konstruktiver Unterstützung und (tendenziell) effizienter Klassenführung mit Inklusion und Bildungsstandards zu verbinden. Bei der Einschätzung der Unterrichtsgestaltung, die mit der Umsetzung beider Reformen einhergeht, scheinen zudem die Einstellungen zu und Erfahrungen mit Inklusion und Bildungsstandards sowie diesbezügliche Selbstwirksamkeitserwartungen eine Rolle zu spielen. Inwiefern berichtete Unterschiede bezogen auf Unterrichtsgestaltung jedoch tatsächlich Unvereinbarkeiten darstellen, bleibt weitgehend eine offene Frage, da Lehrkräfte sowohl inklusivem als auch standardorientiertem Unterricht insgesamt eine (gemessen an den theoretischen Mittelwerten der eingesetzten Skalen) relativ hohe Unterrichtsqualität beimessen. Bezüglich der *Leistungsbeurteilung* deutet sich an, dass mit der Gleichzeitigkeit von Inklusion und Bildungsstandards ein Spannungsverhältnis verbunden ist, das Lehrkräfte möglicherweise vor gegensätzliche Anforderungen bei der Anwendung von Bezugsnormen stellt. Während mit Inklusion die stärkere Ausprägung der individuellen Bezugsnorm verbunden wird, wird mit Bildungsstandards eher die kriteriale Bezugsnorm assoziiert und positive Einstellungen zu Bildungsstandards gehen nach der Auseinandersetzung mit dieser Reform mit einer geringer ausgeprägten individuellen und einer stärker ausgeprägten sozialen BNO einher. Es wurde jedoch diskutiert, dass wahrgenommene gegensätzliche Implikationen beider Reformen für die Leistungsbeurteilung durch eine Kombination unterschiedlicher Bezugsnormen miteinander in Einklang gebracht werden könnten.

Zudem scheint die mit Bildungsstandards einhergehende Denkweise in objektiven Maßstäben bei der Leistungsbeurteilung mit nicht intendierten Folgen für Schülerinnen und Schüler einherzugehen, die stereotypisierten sozialen Gruppen angehören.

Im Rahmen dieser Arbeit konnten Erkenntnisse dazu gewonnen werden, inwiefern für Lehrkräfte, die als betroffene Akteure beide Reformen umsetzen müssen, potentiell gegensätzliche Anforderungen mit Inklusion und Bildungsstandards einhergehen können. Somit konnte der Diskurs um die Vereinbarkeit beider Reformen in Bezug auf Unterrichtsgestaltung und Leistungsbeurteilung einer differenzierten Betrachtung unterzogen werden. Zudem lassen sich aus den Ergebnissen vielfältige Ansatzpunkte für die pädagogische Praxis und für zukünftige Forschung ableiten. Denn sowohl Praxis als auch Forschung stehen vor der Herausforderung, die in dieser Arbeit aufgezeigten potentiell gegensätzlichen Anforderungen, die aus der Sicht von Lehrkräften mit der Gleichzeitigkeit von Inklusion und Bildungsstandards einhergehen können, konstruktiv zu bearbeiten.

## 6 Literaturverzeichnis

- Ahrbeck, B. (2015). Inklusion – ein unerreichbares Ideal? In R. Braches-Chyrek, C. Fischer, C. Mangione, A. Penczek & S. Rahm (Hrsg.), *Herausforderung Inklusion. Schule - Unterricht - Profession* (S. 27–44). Bamberg: University of Bamberg Press.
- Ahrbeck, B. (2016). *Inklusion. Eine Kritik* (3. aktualisierte Auflage). Stuttgart: Kohlhammer.
- Ainscow, M., Booth, T. & Dyson, A. (2006a). *Improving schools, developing inclusion*. London: Routledge.
- Ainscow, M., Booth, T. & Dyson, A. (2006b). Inclusion and the standards agenda. Negotiating policy pressures in England. *International Journal of Inclusive Education*, 10, 295–308.
- Altrichter, H. (2011). Bildungsstandards und Individualisierung im Unterricht - Zwei Reformpolitiken im Spannungsfeld. In F. Hofmann, D. Martinek & U. Schwantner (Hrsg.), *Binnendifferenzierter Unterricht und Bildungsstandards - (k)ein Widerspruch?* (S. 9–27). Wien: Lit.
- Altrichter, H. & Maag Merki, K. (2010). Steuerung der Entwicklung des Schulwesens. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 15–39). Wiesbaden: Springer VS.
- Altrichter, H. & Posch, P. (2007). Analyse erster Erfahrungen mit der Implementation von Bildungsstandards. *Erziehung und Unterricht*, 157, 654–671.
- Amrhein, B. (2014). Am und im Widerspruch arbeiten. Wege aus dem professionellen Unbehagen in inklusiven Bildungsreformen. *Friedrich Jahresheft*, 17–19.
- Artiles, A. J., Kozleski, E. B., Dorn, S. & Christensen, C. (2006). Learning in inclusive education research: re-mediating theory and methods with a transformative agenda. *Review of Research in Education*, 30, 65–108.
- Avramidis, E. & Norwich, B. (2002). Teachers' attitudes towards integration / inclusion. A review of the literature. *European Journal of Special Needs Education*, 17, 129–147.
- Bargh, J. A., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, 62, 893–912.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9, 469–520.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A. et al. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47, 133–180.
- Beauftragte der Bundesregierung für die Belange von Menschen mit Behinderungen (Hrsg.). (2017). *Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderung. Amtliche, gemeinsame Übersetzung von Deutschland, Österreich, Schweiz und Lichtenstein. Stand: Januar 2017*.

- Beer, R. (2006). *Bildungsstandards. Einstellungen von Lehrerinnen und Lehrern*. Wien: Lit.
- Benzing, M., Brändle, T., Klitsche, S., Musekamp, F., Lücken, M. & Thonke, F. (2016). *Hinweise und Anregungen zur Nutzung von KERMIT für die Unterrichts- und Schulentwicklung*. Hamburg: Institut für Bildungsmonitoring und Qualitätsentwicklung.
- BGBI. 2008, Teil II, Nr. 35. *Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen sowie zu dem Fakultativprotokoll vom 13. Dezember 2006 zum Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderungen*. Berlin: Bundesanzeiger.
- Biernat, M. (1995). The shifting standards model: Implications of stereotype accuracy for social judgment. In Y.-T. Lee, L. J. Jussim & C. R. McCauley (Hrsg.), *Stereotype accuracy: Toward appreciating group differences* (S. 87–114). Washington: American Psychological Association.
- Biernat, M. (2009). Stereotypes and Shifting Standards. In T. D. Nelson (Hrsg.), *Handbook of prejudice, stereotyping, and discrimination* (S. 137–152). New York: Psychology Press.
- Biernat, M. & Fuegen, K. (2001). Shifting standards and the evaluation of competence. Complexity in gender-based judgment and decision making. *Journal of Social Issues*, 57, 707–724.
- Biernat, M., Fuegen, K. & Kobrynowicz, D. (2010). Shifting standards and the inference of incompetence. Effects of formal and informal evaluation tools. *Personality and Social Psychology Bulletin*, 36, 855–868.
- Biernat, M. & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5–20.
- Biernat, M., Manis, M. & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60, 485–499.
- Biernat, M. & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched. Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin*, 28, 66–77.
- Biewer, G. (2012). Die neue Welt der Bildungsstandards und ihre erziehungswissenschaftliche Rezeption aus der Perspektive einer Inklusiven Pädagogik. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 81, 9–21.
- Blanck, J. M. (2015). Schulische Integration und Inklusion in Deutschland. Eine vergleichende Betrachtung der 16 Bundesländer. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 153–177). Wiesbaden: Springer VS.
- Blömeke, S., Herzig, B. & Tulodziecki, G. (2007). *Gestaltung von Schule. Eine Einführung in Schultheorie und Schulentwicklung*. Bad Heilbrunn: Klinkhardt.

- Blum, W. (2012). Einführung. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen* (S. 14–32). Berlin: Cornelsen Scriptor.
- Blum, W., Drüke-Noe, C., Hartung, R. & Köller, O. (Hrsg.). (2012). *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Berlin: Cornelsen Scriptor.
- Boban, I. & Hinz, A. (2003). *Index für Inklusion. Lernen und Teilhabe in der Schule der Vielfalt entwickeln*. Zugriff am 17.10.2017. Verfügbar unter <http://www.eenet.org.uk/resources/docs/Index%20German.pdf>
- Boban, I. & Hinz, A. (2013). Auf dem Weg zur inklusiven Schule – mit Hilfe des Index für Inklusion. In V. Moser (Hrsg.), *Die inklusive Schule. Standards für die Umsetzung* (S. 73–78). Stuttgart: Kohlhammer.
- Böhme, K., Kohrt, P., Haag, N. & Weirich, S. (2017). Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 369–385). Münster: Waxmann.
- Bosse, S. & Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empirische Sonderpädagogik*, 4, 279–299.
- Bräu, K. (2008). Die Betreuung selbstständigen Lernens — vom Umgang mit Antinomien und Dilemmata. In G. Breidenstein & F. Schütze (Hrsg.), *Paradoxien in der Reform der Schule* (S. 179–199). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Brokamp, B. (2005). Welche Chancen bieten Bildungsstandards für die Weiterentwicklung der Inklusion? In U. Geiling & A. Hinz (Hrsg.), *Integrationspädagogik im Diskurs. Auf dem Weg zu einer inklusiven Pädagogik?* (S. 113–116). Bad Heilbrunn: Klinkhardt.
- Brügelmann, H. (2015). *Vermessene Schulen - standardisierte Schüler. Zu Risiken und Nebenwirkungen von PISA, Hattie, VerA & Co*. Weinheim: Beltz.
- Brüggemann, C. & Tegge, D. (2016). Kennziffern in der Kritik: Anmerkungen zur indikatoren gestützten Darstellung von Inklusion in der internationalen Bildungsberichterstattung. *Zeitschrift für Inklusion* (1).
- Brunner, M., Emmrich, R., Gärtner, H., Penk, C., Vettorazzi, K. & Wendt, W. (2016). *ISQ-Bericht zur Schulqualität 2016*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Budde, J. & Blasse, N. (2017). Forschung zu inklusivem Unterricht. In B. Lütje-Klose, S. Miller, S. Schwab & B. Streese (Hrsg.), *Inklusion: Profile für die Schul- und Unterrichtsentwicklung in Deutschland, Österreich und der Schweiz. Theoretische Grundlagen - Empirische Befunde - Praxisbeispiele* (S. 239–252). Münster: Waxmann.

- Büker, P. & Meier, S. (2017). Wie kann ein inklusionsgerechtes Leistungsverständnis vom Kind aus gedacht werden? In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung* (S. 165–172). Bad Heilbrunn: Klinkhardt.
- Buschkühle, C.-P., Duncker, L. & Oswalt, V. (2009). Vorwort der Herausgeber. In C.-P. Buschkühle, L. Duncker & V. Oswalt (Hrsg.), *Bildung zwischen Standardisierung und Heterogenität. Ein interdisziplinärer Diskurs* (S. 7–9). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Carle, U. (2017). Eckpunkte für die Entwicklung inklusiven Unterrichts. In F. Hellmich & E. Blumber (Hrsg.), *Inklusiver Unterricht in der Grundschule* (S. 15–33). Stuttgart: Kohlhammer.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Münster: Waxmann.
- Cornett-Ruiz, S. & Hendricks, B. (1993). Effects of Labeling and ADHD Behaviors on Peer and Teacher Judgments. *The Journal of Educational Research*, 86, 349–355.
- Davies, M. & Elliott, S. N. (2012). Inclusive Assessment and Accountability. Policy to evidence-based practices. *International Journal of Disability, Development and Education*, 59, 1–6.
- Davis, P. & Florian, L. (2004). *Teaching strategies and approaches for pupils with special educational needs: A scoping study. DfES Research Report*. London: Department for Education and Skills.
- de Boer, A. A., Pijl, S. J. & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education. A review of the literature. *International Journal of Inclusive Education*, 15, 331–353.
- Dickhäuser, O., Janke, S., Praetorius, A.-K. & Dresel, M. (2017). The Effects of Teachers' Reference Norm Orientations on Students' Implicit Theories and Academic Self-Concepts. *Zeitschrift für Pädagogische Psychologie*, 31, 205–219.
- Dickhäuser, O. & Rheinberg, F. (2003). Bezugsnormorientierung: Erfassung, Probleme, Perspektiven. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Selbstkonzept und Motivation und Selbstregulation* (S. 41–55). Göttingen: Hogrefe.
- Drieschner, E. (2009). *Bildungsstandards praktisch*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Drüke-Noe, C. (2012). Projektorientierung. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen* (S. 126–134). Berlin: Cornelsen Scriptor.
- Dyson, A., Gallannaugh, F. & Millward, A. (2003). Making Space in the Standards Agenda: developing inclusive practices in schools. *European Educational Research Journal*, 2, 228–244.
- EASIE = European Agency for Development in Special Needs Education (Hrsg.). (2017). *European Agency Statistics on Inclusive Education: 2014 Dataset Cross-Country Report*. Odense: European Agency for Development in Special Needs Education.



- Eichfeld, C. & Algermissen, P. (2016). Aber nicht in Sachsen...? Synopse von Studien zu Einstellungen sächsischer Grundschullehrkräfte zu inklusivem Unterricht. In K. Liebers, B. Landwehr, S. Reinhold, S. Riegler & R. Schmidt (Hrsg.), *Facetten grundschulpädagogischer und -didaktischer Forschung* (S. 67–72). Wiesbaden: Springer VS.
- Emmer, E. T. & Stough, L. M. (2001). Classroom Management. A Critical Part of Educational Psychology, With Implications for Teacher Education. *Educational Psychologist*, 36, 103–112.
- Emmrich, R., Ernst, C.-M., Harych, P. & Meiering, G. (2012). *Vergleichsarbeiten der Jahrgangsstufe 3 in Berlin als Beitrag zur Schul- und Unterrichtsentwicklung*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Emmrich, R., Ernst, C.-M., Harych, P. & Wesselhöft, K. (2012). *Vergleichsarbeiten der Jahrgangsstufe 8 in Berlin als Beitrag zur Schul- und Unterrichtsentwicklung*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Eversheim, U. (2015). Bildungsstandards versus Inklusion? - Probleme und Potentiale der Standard- und Kompetenzorientierung für einen inklusiven (Sport)Unterricht. In S. Meier (Hrsg.), *Inklusion als Herausforderung. Aufgabe und Chance für den Schulsport* (S. 207–224). Berlin: Logos.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school. Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.
- Fazio, R. H., Chen, J., McDonel, E. C. & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, 18, 339–357.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C. & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–238.
- Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, 7, 117–140.
- Fiske, S. T. & Neuberg, S. L. (1990). A continuum of impression formation, from category based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74.
- Foster, G. G., Schmidt, C. R. & Sabatino, D. (1976). Teacher Expectancies and the Label “Learning Disabilities”. *Journal of Learning Disabilities*, 9, 111–114.
- Gasterstädt, J. & Urban, M. (2016). Einstellung zu Inklusion? Implikationen aus Sicht qualitativer Forschung im Kontext der Entwicklung inklusiver Schulen. *Empirische Sonderpädagogik*, 54–66.
- Gebhardt, M., Schwab, S., Krammer, M., Gasteiger-Klicpera, B. & Sälzer, C. (2014). Erfassung von individualisiertem Unterricht in der Sekundarstufe I. Eine Quantitative Überprüfung der Skala „Individualisierter Unterricht“ in zwei Schuluntersuchungen in der Steiermark. *Zeitschrift für Bildungsforschung*, 4, 303–316.

- Gebhardt, M., Schwab, S., Nusser, L. & Hessels, M. G. P. (2015). Einstellungen und Selbstwirksamkeit von Lehrerinnen und Lehrern zur schulischen Inklusion in Deutschland - eine Analyse mit Daten des Nationalen Bildungspanels Deutschlands (NEPS). *Empirische Pädagogik*, 29, 211–229.
- Geiling, U. & Gille, N. (2005). Integrative/ inklusive Qualität, Bildungsstandards und Bildungsbarrieren. In U. Geiling & A. Hinz (Hrsg.), *Integrationspädagogik im Diskurs. Auf dem Weg zu einer inklusiven Pädagogik?* (S. 103–106). Bad Heilbrunn: Klinkhardt.
- Gläser-Zikuda, M. (2010). Leistungsvoraussetzungen diagnostizieren und Fördermaßnahmen realisieren. In T. Bohl, W. Helsper, H. G. Holtappels & C. Schelle (Hrsg.), *Handbuch Schulentwicklung* (S. 369–376). Bad Heilbrunn: Klinkhardt.
- Glasman, L. R. & Albarracín, D. (2006). Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. *Psychological bulletin*, 132, 778–822.
- Glock, S. & Krolak-Schwerdt, S. (2013). Does nationality matter? The impact of stereotypical expectations on student teachers' judgments. *Social Psychology of Education*, 16, 111–127.
- Göransson, N. & Niholm, C. (2014). Conceptual diversities and empirical shortcomings – a critical analysis of research on inclusive education. *European Journal of Special Needs Education*, 19, 265-280.
- Greiten, S. (2017). „Leistung“ wird relativ. Konzeptveränderungen zur „Leistung“ bei Lehrkräften durch den Unterricht mit integrativen Lerngruppen. In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung* (S. 157–164). Bad Heilbrunn: Klinkhardt.
- Gresch, C., Rjosk, C., Kocaj, A. & Stanat, P. (2017). Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Deutschland. Einleitung. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 277–281). Münster: Waxmann.
- Griffith, G. & Scharmann, L. (2008). Initial impacts of No Child Left Behind on elementary science education. *Journal of Elementary Science Education*, 20 (3), 35–48.
- Grillitsch, M. (2010). *Bildungsstandards auf dem Weg in die Praxis. Ergebnisse einer Befragung von Lehrkräften und Schulleiter/innen der Sekundarstufe I zur Rezeption der Bildungsstandards und deren Implementation*. Graz: Leykam.
- Grillitsch, M. & Amtmann, E. (2012). *Rezeption der Bildungsstandards und Erfahrungen mit der Ergebnisrückmeldung. Befunde der Begleitforschung zur Baseline-Testung auf der 4. Schulstufe*. Graz: BIFIE.
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.

- Groß Ophoff, J. (2013). *Lernstandserhebungen: Reflexion und Nutzung*. Münster: Waxmann.
- Haeberlin, U. (1991). Die Integration von leistungsschwachen Schülern. Ein Überblick über empirische Forschungsergebnisse zu Wirkungen von Regelklassen, Integrationsklassen und Sonderklassen auf „Lernbehinderte“. *Zeitschrift für Pädagogik*, 37, 167–189.
- Halbheer, U. & Reusser, K. (2008). Outputsteuerung, Accountability, Educational Governance. Einführung in Geschichte, Begrifflichkeiten und Funktionen von Bildungsstandards. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 26, 253–266.
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J. et al. (2007). *Standards-based accountability under no child left behind. Experiences of teachers and administrators in three states*. Santa Monica, Californien: Rand.
- Harych, P., Emmrich, R. & Pant, H. A. (2017). *Kompetenzstufenmodelle und praktische Unterrichtsentwicklung. Workshop im Rahmen der Tagung „Bildungsstandards im Spannungsfeld zwischen Politik und schulischem Alltag“*, Linz. Johannes Kepler Universität. Zugriff am 27.11.2017. Verfügbar unter [http://paedpsych.jku.at/dev/wp-content/uploads/2016/01/W1\\_Harych.pdf](http://paedpsych.jku.at/dev/wp-content/uploads/2016/01/W1_Harych.pdf)
- Hecht, P., Niedermair, C. & Feyerer, E. (2016). Einstellungen und inklusionsbezogene Selbstwirksamkeitsüberzeugungen von Lehramtsstudierenden und Lehrpersonen im Berufseinstieg – Messverfahren und Befunde aus einem Mixed-Methods-Design. *Empirische Sonderpädagogik*, 86–102.
- Heimlich, U. (2007). Didaktik des gemeinsamen Unterrichts. In J. Walter & F. Wember (Hrsg.), *Sonderpädagogik des Lernens* (S. 357–374). Göttingen: Hogrefe.
- Heinrich, M. (2010). Zum Problem der Anerkennung fragiler Bildungsprozesse innerhalb neuer Steuerung und demokratischer Governance. In S. Aufenanger, F. Hamburger, R. Ludwit & R. Tippelt (Hrsg.), *Bildung in der Demokratie. Beiträge zum 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft* (S. 125–143). Opladen: Budrich.
- Heinzel, F. & Prengel, A. (2012). Heterogenität als Grundbegriff inklusiver Pädagogik. *Zeitschrift für Inklusion* (3). Verfügbar unter <https://www.inklusion-online.net/index.php/inklusion-online/article/view/39/39>
- Hellmich, F. & Görel, G. (2014). Erklärungsfaktoren für Einstellungen von Lehrerinnen und Lehrern zum inklusiven Unterricht in der Grundschule. *Zeitschrift für Bildungsforschung*, 4 (3), 227–240.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer.
- Helmke, A. & Hosenfeld, I. (2003). Vergleichsarbeiten (VERA): eine Standortbestimmung zur Sicherung schulischer Kompetenzen - Teil 1: Grundlagen, Ziele, Realisierung. *Schulverwaltung NRW, Ausgabe Hessen/Rheinland-Pfalz/Saarland* (1), 10–13.
- Helsper, W. (2002). Lehrerprofessionalität als antinomische Handlungsstruktur. In M. Kraul, W. Marotzki & C. Schweppe (Hrsg.), *Biographie und Profession* (S. 64–102). Bad Heilbrunn: Klinkhardt.

- Henke, T., Koch, H. & Spörer, N. (2015). Einschätzungen von Schülerinnen und Schülern zur Unterrichtsgestaltung. In N. Spörer, A. Schröder-Lenzen, M. Vock & K. Maaz (Hrsg.), *Inklusives Lernen und Lehren im Land Brandenburg. Abschlussbericht zur Begleitforschung des Pilotprojekts „Inklusive Grundschule“* (S. 181–196). Ludwigfelde, Potsdam: Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM), Universität Potsdam.
- Hennies, J. & Ritter, M. (Hrsg.). (2015). *Deutschunterricht in der Inklusion. Auf dem Weg zu einer inklusiven Deutschdidaktik*. Stuttgart: Klett.
- Hillenbrand, C. (2011). *Didaktik bei Unterrichts- und Verhaltensstörungen* (3. aktualisierte Auflage). München: Ernst Reinhardt.
- Hilton, J. L. & von Hippel, W. (1996). Stereotypes. *Annual review of psychology*, 47, 237–271.
- Hinz, A. (2002). Von der Integration zur Inklusion – terminologisches Spiel oder konzeptionelle Weiterentwicklung. *Zeitschrift für Heilpädagogik*, 53, 354–361.
- Hinz, A. (2003). Die Debatte um Integration und Inklusion – Grundlage für aktuelle Kontroversen in Behindertenpolitik und Sonderpädagogik? *Sonderpädagogische Förderung*, 48, 330–347.
- Hinz, A. (2009). Inklusive Pädagogik in der Schule - veränderter Orientierungsrahmen für die schulische Sonderpädagogik!? Oder doch deren Ende?? *Zeitschrift für Heilpädagogik*, 60, 171–179.
- Hinz, A. (2010). Aktuelle Erträge der Debatte um Inklusion – worin besteht der „Mehrwert“ gegenüber Integration? In Evangelische Stiftung Alsterdorf & Katholische Hochschule für Sozialwesen Berlin (Hrsg.), *Enabling Community. Anstöße für Politik und soziale Praxis* (S. 191–202). Hamburg: Alsterdorf Verlag.
- Hinz, A. (2013). Inklusion – von der Unkenntnis zur Unkenntlichkeit!? Kritische Anmerkungen zu einem Jahrzehnt Diskurs über schulische Inklusion in Deutschland. *Zeitschrift für Inklusion* (1).
- Hirschauer, M. & Kullmann, H. (2010). Lehrerprofessionalität im Zeichen von Heterogenität – Stereotype bei Lehrkräften als kollegial zu bearbeitende Herausforderung. In J. Hagedorn, Schurt V., C. Steber & W. Waburg (Hrsg.), *Ethnizität, Geschlecht, Familie und Schule. Heterogenität als erziehungswissenschaftliche Herausforderung* (S. 351–373). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hofmann, F. (2011). Bildungsstandards und Individualisierung - ein Widerspruch? Unterschiedliche Reaktionsmuster von Lehrpersonen - ein Hypothesen generierender Beitrag. In F. Hofmann, D. Martinek & U. Schwantner (Hrsg.), *Binnendifferenzierter Unterricht und Bildungsstandards - (k)ein Widerspruch?* (S. 47–62). Wien: Lit.
- Holmeier, M. (2012). Bezugsnormorientierung im Unterricht im Kontext zentraler Abiturprüfungen. In K. Maag Merki (Hrsg.), *Zentralabitur* (S. 237–261). Wiesbaden: Springer VS.
- Holmeier, M. (2013). *Leistungsbeurteilung im Zentralabitur*. Wiesbaden: Springer VS.

- Hörmann, B. & Hopmann, S. T. (2011). Marginalisierungsprozesse von Menschen mit Behinderungen im Rahmen von „School-Accountability“-Maßnahmen. In P. Flieger & V. Schönwiese (Hrsg.), *Menschenrechte – Integration – Inklusion. Aktuelle Perspektiven aus der Forschung* (S. 105–110). Bad Heilbrunn: Klinkhardt.
- Institut für Qualitätsentwicklung Mecklenburg-Vorpommern (Hrsg.). (2014). *Handreichung zur Benotung von Teilen der Vergleichsarbeiten (VerA) als schriftliche Lernerfolgskontrolle*. Zugriff am 25.10.2017. Verfügbar unter [http://www.bildung-mv.de/export/sites/bildungsserver/downloads/Handreichung\\_VERA\\_05\\_2014.pdf](http://www.bildung-mv.de/export/sites/bildungsserver/downloads/Handreichung_VERA_05_2014.pdf)
- Jantowski, A. (2014). Beurteilung von Schülerleistungen im Spannungsfeld von Standardorientierung und individueller Förderung. In A. Jantowski & R. Möllers (Hrsg.), *Unterricht im Spannungsfeld zwischen Kompetenz- und Standardorientierung* (S. 26–34). Bad Berka: Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien.
- Jordan, A., Schwartz, E. & McGhie-Richmond, D. (2009). Preparing teachers for inclusive classrooms. *Teaching and Teacher Education*, 25, 535-542.
- Jordan, A. & Stanovich, P. (2001). Patterns of Teacher-Student Interaction in Inclusive Elementary Classrooms and Correlates with Student Self-Concept. *International Journal of Disability, Development and Education*, 48, 33–52.
- Katzenbach, D. (2010). Bildung und Anerkennung. In O. Musenberg & J. Riegert (Hrsg.), *Bildung und geistige Behinderung: Bildungstheoretische Reflexionen* (S. 93–114). Oberhausen: Athena.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S. & Russin, A. (2000). Just say no (to stereotyping): effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, 78, 871–888.
- Kessels, U., Erbring, S. & Heiermann, L. (2014). Implizite Einstellungen von Lehramtsstudierenden zur Inklusion. *Psychologie in Erziehung und Unterricht*, 61, 189–202.
- Klein, H. P. (2010). Die neue Kompetenzorientierung: Exzellenz oder Nivellierung? *Journal für Didaktik der Biowissenschaften*, 1, 1–11.
- Klemm, K. (2015). *Inklusion in Deutschland. Daten und Fakten*. Gütersloh: Bertelsmann Stiftung.
- Klieme, E. (2003). Benotungsmaßstäbe an Schulen: Pädagogische Praxis und institutionelle Bedingungen. Eine empirische Analyse auf der Basis der PISA-Studie. In H. Döbert, B. von Kopp, R. Martini & M. Weiß (Hrsg.), *Bildung vor neuen Herausforderungen* (S. 195–210). Neuwied: Luchterhand.
- Klieme, E. (2006). Empirische Unterrichtsforschung: aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52, 765–773.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung.

- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52, 876–903.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung im internationalen Vergleich. In E. Klieme & J. Baumert (Hrsg.), *TIMSS -Impulse für Schule und Unterricht* (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.
- Klieme, E. & Warwas, J. (2011). Konzepte der Individuellen Förderung. *Zeitschrift für Pädagogik*, 57, 805–818.
- KMK (Hrsg.). (1994). *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluss der Kultusministerkonferenz vom 06.05.1994.* Zugriff am 15.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/1994/1994\\_05\\_06-Empfehl-Sonderpaedagogische-Foerderung.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1994/1994_05_06-Empfehl-Sonderpaedagogische-Foerderung.pdf)
- KMK (Hrsg.). (2003a). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 04.12.2003.* München: Wolters Kluwer.
- KMK (Hrsg.). (2003b). *Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 04.12.2003.* Köln: Wolters Kluwer.
- KMK (Hrsg.). (2003c). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 04.12.2003.* Köln: Wolters Kluwer.
- KMK (Hrsg.). (2003d). *Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 04.12.2003.* Zugriff am 25.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Bildungsstandards-Mittleren-SA-Bio-Che-Phy.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Mittleren-SA-Bio-Che-Phy.pdf)
- KMK (Hrsg.). (2004a). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004b). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Hauptschulabschluss (Jahrgangsstufe 9). Beschluss der Kultusministerkonferenz vom 15.10.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004c). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 16.12.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004d). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 16.12.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004e). *Bildungsstandards im Fach Deutsch für den Hauptschulabschluss (Jahrgangsstufe 9). Beschluss der Kultusministerkonferenz vom 15.10.2004.* München/Neuwied: Wolters Kluwer.

- KMK (Hrsg.). (2004f). *Bildungsstandards im Fach Deutsch für den Primarbereich (Jahrgangsstufe 4). Beschluss der Kultusministerkonferenz vom 15.10.2004.* München: Wolters Kluwer.
- KMK (Hrsg.). (2004g). *Bildungsstandards im Fach Mathematik für den Hauptschulabschluss (Jahrgangsstufe 9). Beschluss der Kultusministerkonferenz vom 15.10.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004h). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4). Beschluss der Kultusministerkonferenz vom 15.10.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004i). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss der Kultusministerkonferenz vom 16.12.2004.* München/Neuwied: Wolters Kluwer.
- KMK (Hrsg.). (2004j). *Vereinbarung über Bildungsstandards für den Hauptschulabschluss (Jahrgangsstufe 9). Beschluss der Kultusministerkonferenz vom 15.10.2004.* Zugriff am 25.10.2017. Verfügbar unter [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2004/2004\\_10\\_15-Bildungsstandards-Haupt.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Haupt.pdf)
- KMK (Hrsg.). (2004k). *Vereinbarung über Bildungsstandards für den Primarbereich (Jahrgangsstufe 4). Beschluss der Kultusministerkonferenz vom 15.10.2004.* Zugriff am 25.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2004/2004\\_10\\_15-Bildungsstandards-Primar.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Primar.pdf)
- KMK (Hrsg.). (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung.* Zugriff am 16.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2010/2010\\_00\\_00-Konzeption-Bildungsstandards.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2010/2010_00_00-Konzeption-Bildungsstandards.pdf)
- KMK (Hrsg.). (2011). *Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen. Beschluss der Kultusministerkonferenz vom 20.10.2011.* Zugriff am 09.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2011/2011\\_10\\_20-Inklusive-Bildung.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf)
- KMK (Hrsg.). (2012a). *Bildungsstandards für die fortgeführte Fremdsprache (Englisch/Französisch) für die Allgemeine Hochschulreife. Beschluss der Kultusministerkonferenz vom 18.10.2012.* Zugriff am 09.10.2017. Verfügbar unter [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2012/2012\\_10\\_18-Bildungsstandards-Deutsch-Abi.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Deutsch-Abi.pdf)
- KMK (Hrsg.). (2012b). *Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife. Beschluss der Kultusministerkonferenz vom 18.10.2012.* Zugriff am 09.10.2017. Verfügbar unter [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2012/2012\\_10\\_18-Bildungsstandards-Deutsch-Abi.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Deutsch-Abi.pdf)
- KMK (Hrsg.). (2012c). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife. Beschluss der Kultusministerkonferenz vom 18.10.2012.* Zugriff am 09.10.2017. Verfügbar unter [https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2012/2012\\_10\\_18-Bildungsstandards-Deutsch-Abi.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Deutsch-Abi.pdf)

- KMK (Hrsg.). (2012d). *Vereinbarung zur Weiterentwicklung von VERA. Beschluss der Kultusministerkonferenz vom 08.03.2012*. Zugriff am 12.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2012/2012\\_03\\_08\\_Weiterentwicklung-VERA.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_03_08_Weiterentwicklung-VERA.pdf)
- KMK (Hrsg.). (2013). *Kompetenzstufenmodell zu den Bildungsstandards für das Fach Deutsch im Kompetenzbereich „Sprechen und Zuhören“. Primarbereich. Beschluss der Kultusministerkonferenz vom 04.03.2010. Auf Grundlage des Ländervergleichs 2011 überarbeiteter Entwurf in der Version vom 13.02.2013*. Zugriff am 15.10.2017. Verfügbar unter [https://www.iqb.hu-berlin.de/bista/ksm/KSM\\_GS\\_Deutsch\\_Z\\_2.pdf](https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Deutsch_Z_2.pdf)
- KMK (Hrsg.). (2016a). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Zugriff am 09.10.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2015/2015\\_06\\_11-Gesamtstrategie-Bildungsmonitoring.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf)
- KMK (Hrsg.). (2016b). *Sonderpädagogische Förderung in allgemeinen Schulen (ohne Förderschulen) 2015/2016*. Zugriff am 09.10.2017. Verfügbar unter [https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus\\_SoPa\\_e\\_Int\\_2015.pdf](https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPa_e_Int_2015.pdf)
- KMK (Hrsg.). (2017). *Verfahren zur Qualitätssicherung auf Schulebene*. Zugriff am 10.10.2017. Verfügbar unter <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring/verfahren-zur-qualitaetssicherung-auf-schulebene.html>
- Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A. & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 212–233). Weinheim: Beltz-Juventa.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191.
- Köller, O. (2005). Bezugsnormorientierung von Lehrkräften: Konzeptuelle Grundlagen, empirische Befunde und Ratschläge für praktisches Handeln. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendungen* (S. 189–202). Stuttgart: Kohlhammer.
- Köller, O. (2009a). Bildungsstandards in Deutschland: Implikationen für die Qualitätssicherung und Unterrichtsqualität. In M. A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik. 9. Sonderheft der Zeitschrift für Erziehungswissenschaft* (S. 47–59). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O. (2009b). Die institutionelle Umsetzung durch das IQB – Konzeption und Kritik. In R. Wernstedt & M. John-Ohnesorg (Hrsg.), *Bildungsstandards als Instrument schulischer Qualitätsentwicklung. Zementierung des Selektionsprinzips oder Mittel zur Chancengerechtigkeit?* (S. 28–34). Berlin: Friedrich-Ebert-Stiftung.



- Köller, O. (2010). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 529–548). Wiesbaden: Springer VS.
- Kopcha, T. J. & Sullivan, H. (2007). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development*, 55, 627–646.
- Kottmann, B. & Miller, S. (2014). Grundschullehrkräfte im Entscheidungsdilemma zwischen Fördern und Selektieren. In B. Kopp, S. Martschinke, M. Munser-Kiefer, M. Haider, E.-M. Kirschhock, G. Ranger et al. (Hrsg.), *Individuelle Förderung und Lernen in der Gemeinschaft* (S. 218–221). Wiesbaden: Springer VS.
- Kühle, B. & Peek, R. (2007). Lernstandserhebungen in Nordrhein-Westfalen: Evaluationsbefunde zur Rezeption und zum Umgang mit Ergebnisrückmeldungen in Schulen. *Empirische Pädagogik*, 21, 428–447.
- Kühn, S. M. & Racherbäumer, K. (2013). Standardisierung und/oder Individualisierung? Empirische Befunde zur Umsetzung von Maßnahmen zur individuellen Förderung im Kontext zentraler Abschlussprüfungen. *Unterrichtswissenschaft*, 41, 173–190.
- Kullmann, H., Lütje-Klose, B. & Textor, A. (2014). Allgemeine Didaktik für inklusive Lerngruppen – fünf Leitprinzipien als Grundlage eines Bielefelder Ansatzes der inklusiven Didaktik. In B. Amrhein & M. Dziak-Mahler (Hrsg.), *Fachdidaktik inklusiv. Auf der Suche nach didaktischen Leitlinien für den Umgang mit Vielfalt in der Schule* (S. 89–107). Münster: Waxmann.
- Kunter, M. & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–32). Münster: Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers. Effects on instructional quality and student development. *Journal of Educational Psychology*, 105, 805–820.
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: UTB.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. (S. 85–113). Münster: Waxmann.
- Lange, V. (2017). *Inklusive Bildung in Deutschland. Ländervergleich*. Berlin: Friedrich-Ebert-Stiftung.
- Lankes, E. M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E. M. Lankes et al. (Hrsg.), *IGLU 2006 Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.
- Lenski, A., Richter, D. & Pant, H. (2015). Kompetenzorientierung im Unterricht aus der Perspektive von Lehrkräften und Schülerinnen und Schülern. *Zeitschrift für Pädagogik*, 61, 712–737.

- Leutner, D., Fleischer, J., Spoden, C. & Wirth, J. (2008). Landesweite Lernstandserhebungen zwischen Bildungsmonitoring und Individualdiagnostik. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 149–167). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Leyens, J. P., Yzerbyt, V. & Schadron, G. (1994). *Stereotypes and social cognition*. London: Sage Publications.
- Li, Q. (1999). Teachers' beliefs and gender differences in mathematics. A review. *Educational Research*, 41, 63–76.
- Liebers, K. & Schmidt, C. (2017). Formatives Assessment im inklusiven Unterricht – Forschungsstand und erste Befunde. In F. Hellmich & E. Blumber (Hrsg.), *Inklusiver Unterricht in der Grundschule* (S. 50–65). Stuttgart: Kohlhammer.
- Liebers, K., Seifert, C. & Kolke, S. (2014). Gestaltung inklusiven Unterrichts in den ERINA-Oberschulen – erste Befunde. In M. Lichtblau, D. Blömer, A.-K. Jüttner, K. Koch, M. Krüger & R. Werning (Hrsg.), *Forschung zu inklusiver Bildung. Gemeinsam anders lehren und lernen* (S. 230–243). Bad Heilbrunn: Klinkhardt.
- Lindsay, G. (2007). Educational psychology and the effectiveness of inclusive education/mainstreaming. *The British journal of educational psychology*, 77, 1–24.
- Lorenz, G., Gentrup, S., Kristen, C., Stanat, P. & Kogan, I. (2016). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 68, 89–111.
- Löser, J. M. & Werning, R. (2015). Inklusion - allgegenwärtig, kontrovers, diffus? *Erziehungswissenschaft*, 26 (51), 17–24.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish–little-pond effect. *Contemporary Educational Psychology*, 30, 263–285.
- Maag Merki, K. (2010). Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 145–169). Wiesbaden: Springer VS.
- Maier, U. (2008a). Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. *Zeitschrift für Erziehungswissenschaft*, 11, 453–474.
- Maier, U. (2008b). Rezeption und Nutzung von Vergleichsarbeiten aus der Perspektive von Lehrkräften. *Zeitschrift für Pädagogik*, 54, 95–117.
- Maier, U. & Kuper, H. (2012). Vergleichsarbeiten als Instrumente der Qualitätsentwicklung an Schulen. *Die Deutsche Schule*, 104, 88–99.
- Marsh, H. W. (1984). Relations among dimensions of self-attribution, dimensions of self-concept, and academic achievements. *Journal of Educational Psychology*, 76, 1291–1308.
- Marsh, H. W. & Hau, K.-T. (2003). Big-Fish-Little-Pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58, 364–376.

- Meier, S., Haut, J. & Ruin, S. (2016). Leistung als Selbstverständlichkeit? Eine kritische Reflexion divergierender Leistungsverständnisse (im Sport) vor dem Hintergrund des Inklusionsdiskurses. *Zeitschrift für Inklusion* (3).
- Ministerium für Schule und Weiterbildung (Hrsg.). (2008). *Richtlinien und Lehrpläne für die Grundschule in Nordrhein-Westfalen. Deutsch, Sachunterricht, Mathematik, Englisch, Musik, Kunst, Sport, Evangelische Religionslehre, Katholische Religionslehre*. Frechen: Ritterbach Verlag.
- Mischo, C. & Groeben, N. (1995). Bezugsnormorientierung: Warum sich LehrerInnen unterscheiden. *Empirische Pädagogik*, 9, 423–459.
- Mischo, C. & Rheinberg, F. (1995). Erziehungsziele von Lehrern und individuelle Bezugsnormen der Leistungsbewertung. *Zeitschrift für Pädagogische Psychologie*, 9, 139–151.
- Möller, J. (2013). Effekte inklusiver Beschulung aus empirischer Sicht. In J. Baumert, V. Masuhr, J. Möller, T. Riecke-Baulecke, H.-E. Tenorth & R. Werning (Hrsg.), *Inklusion. Forschungsergebnisse und Perspektiven* (Schulmanagement-Handbuch, Bd. 146, S. 15–37). München: Oldenbourg.
- Moser, U. (2003). *Klassenscockpit im Kanton Zürich. Ergebnisse einer Befragung von Lehrerinnen und Lehrern der 6. Klassen über ihre Erfahrungen im Rahmen der Erprobung von Klassenscockpit im Schuljahr 2002/03. Bericht zuhanden der Bildungsdirektion des Kantons Zürich*. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung der Universität Zürich.
- Moser, V., Schäfer, L. & Kropp, A. (2014). Kompetenzbereiche von Lehrkräften in inklusiven Settings. In M. Lichtblau, D. Blömer, A.-K. Jüttner, K. Koch, M. Krüger & R. Werning (Hrsg.), *Forschung zu inklusiver Bildung. Gemeinsam anders lehren und lernen* (S. 124–143). Bad Heilbrunn: Klinkhardt.
- Moser Opitz, E. (2006). PISA und Bildungsstandards: Stein des Anstoßes oder Anstoß für die Sonderpädagogik? *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 75, 110–120.
- Moser Opitz, E. (2014). Inklusive Didaktik im Spannungsfeld von gemeinsamen Lernen und effektiver Förderung. Ein Forschungsüberblick und eine Analyse von didaktischen Konzeptionen für inklusiven Unterricht. In K. Zierer & K. Reusser (Hrsg.), *Jahrbuch für allgemeine Didaktik*. (S. 52–68). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Nachtigall, C. & Hellrung, M. (2013). Zur zeitlichen Entwicklung der Rezeption von Vergleichsarbeiten. *Empirische Pädagogik*, 27, 423–441.
- Nürnberger, M., Nerb, J., Schmitz, F., Keller, J. & Sütterlin, S. (2016). Implicit Gender Stereotypes and Essentialist Beliefs Predict Preservice Teachers' Tracking Recommendations. *The Journal of Experimental Education*, 84, 152–174.
- Nusser, L. & Wolter, I. (2016). There's plenty more fish in the sea. Das akademische Selbstkonzept von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen in integrativen und segregierten Schulsettings. *Empirische Pädagogik*, 30, 130-143.
- Oelkers, J. & Reusser, K. (2008). *Qualität entwickeln, Standards sichern, mit Differenz umgehen*. Bonn und Berlin: Bundesministerium für Bildung und Forschung.

- Oser, F. K. & Baeriswyl, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *AERA's Handbook of Research on Teaching* (S. 1031–1065). Washington: American Educational Research Association.
- Pant, H. A. (2013). Wer hat einen Nutzen von Kompetenzmodellen? *Zeitschrift für Erziehungswissenschaft*, 16, 71–79.
- Pant, H. A. (2016). *Einführung in den Bildungsplan 2016. Warum ein »neuer« Bildungsplan? Anlässe und Absichten der Bildungsplanreform*. Zugriff am 01.11.2017. Verfügbar unter <http://www.bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/EINFUEHRUNG>
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 64, 827–845.
- Plank, S. B. & Falk Condliffe, B. (2013). Pressures of the Season: An Examination of Classroom Quality and High-Stakes Accountability. *American Educational Research Journal*, 50, 152–1182.
- Pool Maag, S. & Moser Opitz, E. (2014). Inklusiver Unterricht – grundsätzliche Fragen und Ergebnisse einer explorativen Studie. *Empirische Sonderpädagogik*, 133–149.
- Prenzel, A. (2015a). Didaktische Diagnostik als Element alltäglicher Lehrarbeit – „Formatives Assessment“ im inklusiven Unterricht. In B. Amrhein & K. Ziemer (Hrsg.), *Diagnostik im Kontext inklusiver Bildung – Theorien, Ambivalenzen, Akteure* (S. 49–63). Bad Heilbrunn: Klinkhardt.
- Prenzel, A. (2015b). Inklusive Bildung: Grundlagen, Praxis, offene Fragen. In T. H. Häcker & M. Walm (Hrsg.), *Inklusion als Entwicklung. Konsequenzen für Schule und Lehrerbildung* (S. 27–46). Bad Heilbrunn: Klinkhardt.
- Prenzel, A. (2017). Individualisierung in der „Caring Community“. Zur inklusiven Verbesserung von Lernleistungen. In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung* (S. 13–27). Bad Heilbrunn: Klinkhardt.
- Racherbäumer, K. & Kühn, S. M. (2013). Zentrale Prüfungen und individuelle Förderung. *Zeitschrift für Bildungsforschung*, 3, 27–45.
- Rauch, D., Mang, J., Härtig, H. & Haag, N. (2016). Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 317–348). Münster: Waxmann.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, 76, 85–97.
- Ravitch, D. (1995). *National standards in American education. A citizen's guide*. Washington, D.C.: Brookings.
- Reich, K. (2014). *Inklusive Didaktik. Bausteine für eine inklusive Schule* (Pädagogik). Weinheim: Beltz.

- Retelsdorf, J. & Günther, C. (2011). Achievement goals for teaching and teachers reference norms: Relations with instructional practices. *Teaching and Teacher Education*, 27, 1111–1119.
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Göttingen: Hogrefe.
- Rheinberg, F. (1998). Bezugsnorm-Orientierung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 39–43). Weinheim: Beltz.
- Richard, F. D., Bond, C. F. & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, 7, 331–363.
- Richter, D., Böhme, K., Becker, M., Pant, H. A. & Stanat, P. (2014). Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten. Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen von Schülerinnen und Schülern. *Zeitschrift für Pädagogik*, 60, 225–244.
- Ritter, M. & Hennies, J. (2015). Inklusive Deutschdidaktik. Konzeptionelle Überlegungen zur Grundlegung. In T. H. Häcker & M. Walm (Hrsg.), *Inklusion als Entwicklung. Konsequenzen für Schule und Lehrerbildung* (S. 253–266). Bad Heilbrunn: Klinkhardt.
- Rothland, M. (2013). Beruf: Lehrer/Lehrerin – Arbeitsplatz: Schule Charakteristika der Arbeitstätigkeit und Bedingungen der Berufssituation. In M. Rothland (Hrsg.), *Belastung und Beanspruchung im Lehrerberuf* (S. 21–39). Wiesbaden: Springer VS.
- Ruberg, C. & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik*, 63, 393–415.
- Ruijs, N. M. & Peetsma, T. T. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4, 67–79.
- Sacher, W. (2009). *Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe* (5., überarb. und erw. Aufl.). Bad Heilbrunn: Klinkhardt.
- Sander, A. (2005). Bildungsstandards und Bildungsbarrieren: Thesen aus Perspektive einer inklusiven Pädagogik. In U. Geiling & A. Hinz (Hrsg.), *Integrationspädagogik im Diskurs. Auf dem Weg zu einer inklusiven Pädagogik?* (S. 110–113). Bad Heilbrunn: Klinkhardt.
- Schiepe-Tiska, A., Schmidtner, S., Müller, K., Heine, J.-H., Neumann, K. & Lüdtke, O. (2016). Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 134–175). Münster: Waxmann.
- Schipolowski, S., Stanat, P. & Rjosk, C. (2017). Kontextinformationen zu den Schulsystemen der Länder im Primarbereich. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 115–128). Münster: Waxmann.

- Schnitzler, C. D., Scheerer-Neumann, G. & Ritter, C. (2008). *ILeA - Individuelle Lernstandsanalysen - Lesen*. Ludwigsfelde: Landesinstitut für Schule und Medien Berlin-Brandenburg.
- Schöne, C., Dickhäuser, O., Spinath, B. & Stiensmeier-Pelster, J. (2004). Zielorientierung und Bezugsnormorientierung. Zum Zusammenhang zweier Konzepte. *Zeitschrift für Pädagogische Psychologie*, 18, 93–99.
- Schuck, K. D. (2014). Individualisierung und Standardisierung in der inklusiven Schule – ein unauflösbarer Widerspruch? *Die Deutsche Schule*, 106, 162–174.
- Schwab, S. & Seifert, S. (2015). Einstellungen von Lehramtsstudierenden und Pädagogikstudierenden zur schulischen Inklusion – Ergebnisse einer quantitativen Untersuchung. *Zeitschrift für Bildungsforschung*, 5, 73–87.
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren. (2007). Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule (HarmoS-Konkordat). Zugriff am 24.10.2017. Verfügbar unter [https://edudoc.ch/record/24711/files/HarmoS\\_d.pdf](https://edudoc.ch/record/24711/files/HarmoS_d.pdf)
- Seidel, T., Prenzel, M., Rimmel, R., Dalehefte, I. M., Herweg, C., Kobarg, M. et al. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, 52, 799–821.
- Seidel, T. & Shavelson, R. J. (2007). Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results. *Review of Educational Research*, 77, 454-499.
- Seifert, A. & Müller-Zastrau, A. (2014). Das Dilemma der Leistungsbewertung in der inklusiven Grundschule. In E.-F. Franz, S. Trumpp & I. Esslinger-Hinz (Hrsg.), *Inklusion Eine Herausforderung für die Grundschulpädagogik* (S. 155–165). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Seitz, S. (2006a). Inklusive Didaktik: Die Frage nach dem ‚Kern der Sache‘. *Zeitschrift für Inklusion* (1).
- Seitz, S. (2006b). Inklusive Didaktik nach PISA. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 75, 192–199.
- Seitz, S. & Platte, A. (2006). Unterricht zwischen Vielfalt und Standardisierung. In A. Platte, S. Seitz & K. Terfloth (Hrsg.), *Inklusive Bildungsprozesse* (S. 122–131). Bad Heilbrunn: Klinkhardt.
- Seitz, S. & Scheidt, K. (2012). Vom Reichtum inklusiven Unterrichts - Sechs Ressourcen zur Weiterentwicklung. *Zeitschrift für Inklusion* (1-2).
- Seiz, J., Decristan, J., Kunter, M. & Baumert, J. (2016). Differenzielle Effekte von Klassenführung und Unterstützung für Schülerinnen und Schüler mit Migrationshintergrund. *Zeitschrift für Pädagogische Psychologie*, 30, 237–249.
- Senatsverwaltung für Bildung, Jugend und Familie & Ministerium für Bildung, Jugend und Sport des Landes Brandenburg (Hrsg.). (2015). *Rahmenlehrplan für die Jahrgangsstufen 1 - 10 der Berliner und Brandenburger Schulen*.

- Sermier Dessemontet, R., Benoit, V. & Bless, G. (2011). Schulische Integration von Kindern mit einer geistigen Behinderung. Untersuchung der Entwicklung der Schulleistungen und der adaptiven Fähigkeiten, der Wirkung auf die Lernentwicklung der Mitschüler sowie der Lehrereinstellungen zur Integration. *Empirische Sonderpädagogik*, 3, 291–307.
- Smith, J. M. & Kovacs, P. E. (2011). The impact of standards-based reform on teachers: the case of 'No Child Left Behind'. *Teachers and Teaching. Theory and Practice*, 17, 201–225.
- Solzbacher, C. (2008). Was denken Lehrerinnen und Lehrer über individuelle Förderung? *Pädagogik*, 60 (3), 38–42.
- Spinner, K. H. (2005). Der standardisierte Schüler. Wider den Wunsch, Heterogenität überwinden zu wollen. *Friedrich Jahresheft*, 23, 88–93.
- Sprietsma, M. (2013). Discrimination in grading. Experimental evidence from primary school teachers. *Empirical Economics*, 45, 523–538.
- Stang, J. & Urhahne, D. (2016). Stabilität, Bezugsnormorientierung und Auswirkungen der Urteilsgenauigkeit. *Zeitschrift für Pädagogische Psychologie*, 30, 251–262.
- Statistisches Bundesamt. (2017). *Bildung und Kultur. Allgemeinbildende Schulen. Schuljahr 2016/2017*. Wiesbaden: Statistisches Bundesamt.
- Streese, B., Schiermeyer-Reichl, I., Meyer, A., Moritz, F. & Wenzel, E. (2017). Inklusiv unterrichten - inklusiv bewerten? Impulse zur ‚inklusive Leistungsbewertung‘ in Schulen der Sekundarstufe. In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung* (S. 121–129). Bad Heilbrunn: Klinkhardt.
- Sturm, T. (2015). Inklusion: Kritik und Herausforderung des schulischen Leistungsprinzips. *Erziehungswissenschaft*, 26 (51), 25–32.
- Südkamp, A., Pohl, S., Hardt, K., Jordan, A. K. & Duchhardt, C. (2015). Kompetenzmessung in den Bereichen Lesen und Mathematik bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 243–272). Wiesbaden: Springer VS.
- Tenenbaum, H. R. & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99, 253–273.
- Teumer, S. (2012). Paradoxe Handlungsanforderungen an LehrerInnen im Spannungsfeld der Neugestaltung des Schulanfangs. *Zeitschrift für Inklusion* (3).
- Textor, A. (2015). *Einführung in die Inklusionspädagogik*. Bad Heilbrunn: UTB.
- Textor, A., Grüter, S., Schiermeyer-Reichl, I. & Streese, B. (2017). Leistung inklusive? Inklusion in der Leistungsgesellschaft. Zur Einführung in Band II. In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung* (S. 11–12). Bad Heilbrunn: Klinkhardt.

- Therrien, W. J., Zaman, M. & Banda, D. R. (2010). How Can Meta-Analyses Guide Practice? A Review of the Learning Disability Research Base. *Remedial and Special Education*, 32, 206–218.
- Thüringer Ministerium für Bildung, Wissenschaft und Kultur (Hrsg.). (2016). *Lehrplan für den Erwerb der allgemeinen Hochschulreife. Deutsch*. Zugriff am 01.11.2017. Verfügbar unter <http://www.schulportal-thueringen.de/web/guest/media/detail?tspi=1394&tspt=%3A%3B%3AbackUrl%3A%3D%3A%2Fmedia%2Fdetail%3Ftspi%3D1298>
- Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, 50, 49–62.
- Trautmann, M. & Wischer, B. (2007). Individuell fördern im Unterricht. Was wissen wir über innere Differenzierung? *Pädagogik*, 59 (12), 44–48.
- Trautmann, M. & Wischer, B. (2011). *Heterogenität in der Schule. Eine kritische Einführung*. Wiesbaden: Springer VS.
- Trudewind, C. & Kohne, W. (1982). Bezugsnormorientierung des Lehrers und Motiventwicklung: Zusammenhänge mit Schulleistung, Intelligenz und Merkmalen der häuslichen Umwelt in der Grundschulzeit. In F. Rheinberg (Hrsg.), *Bezugsnormen zur Schulleistungsbewertung. Analyse und Intervention* (S. 115–142). Düsseldorf: Schwann.
- Trumpa, S., Janz, F., Heyl, V. & Seifried, S. (2014). Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. *Zeitschrift für Bildungsforschung*, 4, 241–256.
- UNESCO (Hrsg.). (1994). *Die Salamanca Erklärung und der Aktionsrahmen zur Pädagogik für besondere Bedürfnisse. Angenommen von der Weltkonferenz „Pädagogik für besondere Bedürfnisse: Zugang und Qualität“ (Salamanca, Spanien, 7. - 10. Juni 1994)*.
- United Nations. (2006). *Convention on the rights of persons with disabilities*. Zugriff am 09.10.2017. Verfügbar unter <http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>
- Vock, M. & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht*. Berlin: Friedrich-Ebert-Stiftung.
- Von Barga, I. (2017). Formen der Leistungsbewertung im inklusiven Alltag. In B. Lütje-Klose, S. Miller, S. Schwab & B. Streese (Hrsg.), *Inklusion: Profile für die Schul- und Unterrichtsentwicklung in Deutschland, Österreich und der Schweiz. Theoretische Grundlagen - Empirische Befunde - Praxisbeispiele* (S. 141–152). Münster: Waxmann.
- Wacker, A. & Kramer, J. (2012). Vergleichsarbeiten in Baden-Württemberg. *Zeitschrift für Erziehungswissenschaft*, 15, 683–706.
- Wacker, A., Rohlf, C. & Kramer, J. (2013). Sind Bildungsstandards Innovationsimpulse für Unterricht und Leistungsbeurteilung? Ein Querschnittvergleich der Einschätzungen von Lehrerinnen und Lehrern zu zwei Messzeitpunkten. *Zeitschrift für Bildungsforschung*, 3, 119–136.



- Walm, M., Schultz, C., Häcker, T. & Moser, V. (2017). Diagnostik und Leistungsbewertung im Dienste des Lernens – Theoretische Perspektiven auf ein inklusives Entwicklungsfeld. In A. Textor, S. Grüter, I. Schiermeyer-Reichl & B. Streese (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Unterricht, Leistungsbewertung und Schulentwicklung*. Bad Heilbrunn: Klinkhardt.
- Warwas, J., Hertel, S. & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht. *Zeitschrift für Pädagogik*, 57, 854–867.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim/Basel: Beltz.
- Werning, R. (2010). Inklusion zwischen Innovation und Überforderung. *Zeitschrift für Heilpädagogik*, 61, 284–291.
- Werning, R. (2011). Inklusive Pädagogik - Eine Herausforderung für die Schulentwicklung. *Lernende Schule*, 55 (14), 4–8.
- Werning, R. (2013). Inklusive Schulentwicklung. In V. Moser (Hrsg.), *Die inklusive Schule. Standards für die Umsetzung* (S. 51–63). Stuttgart: Kohlhammer.
- Werning, R. (2014). Stichwort: Schulische Inklusion. *Zeitschrift für Erziehungswissenschaft*, 17, 601–623.
- Werning, R. & Arndt, A.-K. (2015). Unterrichtsgestaltung und Inklusion. In E. Kiel (Hrsg.), *Inklusion im Sekundarbereich* (S. 53–96). Stuttgart: Kohlhammer.
- Werning, R. & Baumert, J. (2013). Inklusion entwickeln: Leitlinien für Schulentwicklung und Lehrerbildung. In J. Baumert, V. Masuhr, J. Möller, T. Riecke-Baulecke, H.-E. Tenorth & R. Werning (Hrsg.), *Inklusion. Forschungsergebnisse und Perspektiven* (Schulmanagement-Handbuch, Bd. 146, S. 38–57). München: Oldenbourg.
- Wilbert, J. & Gerdes, H. (2009). Die Bezugsnormwahl bei der Bewertung schulischer Leistungen durch angehende Lehrkräfte des Förderschwerpunktes Lernen. *Heilpädagogische Forschung*, 35, 122–132.
- Wiley, C. (2000). A Synthesis of Research on the Causes, Effects, and Reduction Strategies of Teacher Stress. *Journal of Instructional Psychology*, 27, 80–87.
- Wilhelm, M. (2009). *Integration in der Sekundarstufe I und II. Wie die Umsetzung im Fachunterricht gelingt*. Weinheim: Beltz.
- Wocken, H. (2012). Rettet die Sonderschulen? - Rettet die Menschenrechte! Ein Appell zu einem differenzierten Diskurs über Dekategorisierung. *Zeitschrift für Inklusion* (4).
- Wubbels, T., Brekelmans, M. & Hooymayers, H. P. (1992). Do teacher ideals distort the self-reports of their interpersonal behavior? *Teaching and Teacher Education*, 8, 47–58.
- Wurster, S. & Richter, D. (2016). Nutzung von Schülerleistungsdaten aus Vergleichsarbeiten und zentralen Abschlussprüfungen für Unterrichtsentwicklung in Brandenburger Fachkonferenzen. *Journal for educational research online*, 8 (3), 159–183.

## **Erklärung**

Hiermit versichere ich, dass ich die Dissertation „Inklusion und Bildungsstandards – gegensätzliche Anforderungen für Unterricht und Leistungsbeurteilung? Empirische Studien zur Perspektive von Lehrkräften“ selbstständig verfasst habe. Sämtliche Hilfsmittel, die ich verwendet habe, sind angegeben. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, Dezember 2017

Katharina Holder

---

## **Lebenslauf**

Der Lebenslauf ist aus Gründen des Datenschutzes nicht enthalten.